

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
20 June 2002 (20.06.2002)

PCT

(10) International Publication Number  
**WO 02/48337 A2**

(51) International Patent Classification?: C12N 14/47,  
15/12, A01K 67/027, A61K 38/17, 39/395, C07K 16/18,  
C12Q 1/68, G01N 33/50, 33/53

(21) International Application Number: PCT/US01/48517

(22) International Filing Date:  
12 December 2001 (12.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/255,639	13 December 2000 (13.12.2000)	US
60/257,852	21 December 2000 (21.12.2000)	US
60/260,105	5 January 2001 (05.01.2001)	US
60/262,932	18 January 2001 (18.01.2001)	US
60/263,096	18 January 2001 (18.01.2001)	US
60/263,090	19 January 2001 (19.01.2001)	US
60/265,926	2 February 2001 (02.02.2001)	US

(71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): GRIFFIN, Jennifer, A. [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). YAO, Monique, G. [US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US). DUGGAN, Brendan, M. [AU/US]; 243 Buena Vista Avenue #306, Sunnyvale, CA 94086 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). DING, Li [CN/US]; 3353 Alma Street #146, Palo Alto, CA 94306 (US). LAL, Preeti, G. [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). LEE, Ernestine, A. [US/US]; 624 Kains Street, Albany, CA 94706 (US). RAMKUMAR, Jayalaxmi [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). THANGAVELU, Kavitha [US/US]; 1950 Montecito Avenue #23, Mountain View, CA 94043 (US). XU, Yuming [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). LEE, Sally [US/US]; 825 East Evelyn, #425, Sunnyvale, CA 94086 (US). TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). NGUYEN, Danniell, B. [US/US]; 1403 Ridgwood Drive, San Jose, CA 95118 (US). WARREN, Bridget, A. [US/US]; 10130 Parkwood Drive #2, Cupertino, CA 95014

(US). HONCHELL, Cynthia, D. [US/US]; 400 Laurel Street #203, San Carlos, CA 94070 (US). GIETZEN, Kimberly, J. [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). GANDHI, Ameena, R. [US/US]; 705 5th Avenue, San Francisco, CA 94118 (US). ARVIZU, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). WALIA, Narinder, K. [US/US]; 890 Davis Street, #205, San Leandro, CA 94577 (US). LU, Yan [CN/US]; 3885 Corrina Way, Palo Alto, CA 94303 (US). ELLIOTT, Vicki, W. [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). LU, Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). HAFALIA, April, J., A. [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). AZIMZAI, Valda [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). KHAN, Farrah, A. [IN/US]; 9445 Harrison Street, Des Plaines, IL 60016 (US). TRAN, Uyen, K. [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SECRETED PROTEINS

(57) Abstract: The invention provides human secreted proteins (SECP) and polynucleotides which identify and encode SECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of SECP.

WO 02/48337 A2

## SECRETED PROTEINS

### TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of secreted proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

### BACKGROUND OF THE INVENTION

Protein transport and secretion are essential for cellular function. Protein transport is mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Proteins that are retained in the plasma membrane contain one or more transmembrane domains, each comprised of about 20 hydrophobic amino acid residues.

Secreted proteins are generally synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include proteins with important roles in cell-to-cell signaling. Such proteins include transmembrane receptors and cell surface markers, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, enzymes, neuropeptides, vasomediators, cell surface markers, and antigen recognition molecules. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

Cell surface markers include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various

differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A.N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, MPs may be heavily glycosylated and may contain an Arginine-Glycine-Aspartate (RGD) tripeptide motif which may play a role in adhesive interactions. MPs include extracellular proteins such as fibronectin, collagen, galectin, vitronectin and its proteolytic derivative somatomedin B; and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) *Kidney Int.* 51:1413-1417; Sjaastad, M.D. and Nelson, W.J. (1997) *BioEssays* 19:47-55.)

Mucins are highly glycosylated glycoproteins that are the major structural component of the mucus gel. The physiological functions of mucins are cytoprotection, mechanical protection, maintenance of viscosity in secretions, and cellular recognition. MUC6 is a human gastric mucin that is also found in gall bladder, pancreas, seminal vesicles, and female reproductive tract (Toribara, N.W. et al. (1997) *J. Biol. Chem.* 272:16398-16403). The MUC6 gene has been mapped to human chromosome 11 (Toribara, N.W. et al. (1993) *J. Biol. Chem.* 268:5879-5885). Hemomucin is a novel *Drosophila* surface mucin that may be involved in the induction of antibacterial effector molecules (Theopold, U. et al. (1996) *J. Biol. Chem.* 271:12708-12715).

Tuftelins are one of four different enamel matrix proteins that have been identified so far. The other three known enamel matrix proteins are the amelogenins, enamelin and ameloblastin. Assembly of the enamel extracellular matrix from these component proteins is believed to be critical in producing a matrix competent to undergo mineral replacement. (Paine, C.T. et al. (1998) *Connect Tissue Res.* 38:257-267). Tuftelin mRNA has been found to be expressed in human ameloblastoma

tumor, a non-mineralized odontogenic tumor (Deutsch, D. et al. (1998) Connect. Tissue Res. 39:177-184).

Olfactomedin-related proteins are extracellular matrix, secreted glycoproteins with conserved C-terminal motifs. They are expressed in a wide variety of tissues and in broad range of species, from *Caenorhabditis elegans* to *Homo sapiens*. Olfactomedin-related proteins comprise a gene family with at least 5 family members in humans. One of the five, TIGR/myocilin protein, is expressed in the eye and is associated with the pathogenesis of glaucoma (Kulkarni, N.H. et al. (2000) Genet. Res. 76:41-50). Research by Yokoyama et al. (1996) found a 135-amino acid protein, termed AMY, having 96% sequence identity with rat neuronal olfactomedin-related ER localized protein in a neuroblastoma cell line cDNA library, suggesting an essential role for AMY in nerve tissue (Yokoyama, M. et al. (1996) DNA Res. 3:311-320). Neuron-specific olfactomedin-related glycoproteins isolated from rat brain cDNA libraries show strong sequence similarity with olfactomedin. This similarity is suggestive of a matrix-related function of these glycoproteins in neurons and neurosecretory cells (Danielson, P.E. et al. (1994) J. Neurosci. Res. 38:468-478).

Mac-2 binding protein is a 90-kD serum protein (90K), a secreted glycoprotein isolated from both the human breast carcinoma cell line SK-BR-3, and human breast milk. It specifically binds to a human macrophage-associated lectin, Mac-2. Structurally, the mature protein is 567 amino acids in length and is preceded by an 18-amino acid leader. There are 16 cysteines and seven potential N-linked glycosylation sites. The first 106 amino acids represent a domain very similar to an ancient protein superfamily defined by a macrophage scavenger receptor cysteine-rich domain (Koths, K. et al. (1993) J. Biol. Chem. 268:14245-14249). 90K is elevated in the serum of subpopulations of AIDS patients and is expressed at varying levels in primary tumor samples and tumor cell lines. Ullrich et al. (1994) have demonstrated that 90K stimulates host defense systems and can induce interleukin-2 secretion. This immune stimulation is proposed to be a result of oncogenic transformation, viral infection or pathogenic invasion (Ullrich, A. et al. (1994) J. Biol. Chem. 269:18401-18407).

Semaphorins are a large group of axonal guidance molecules consisting of at least 30 different members and are found in vertebrates, invertebrates, and even certain viruses. All semaphorins contain the sema domain which is approximately 500 amino acids in length. Neuropilin, a semaphorin receptor, has been shown to promote neurite outgrowth in vitro. The extracellular region of neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of neuropilin have been suggested to have roles in protein-protein interactions and are thought to be involved in the binding of semaphorins through the sema and the C-terminal domains (reviewed in Raper, J.A. (2000) Curr. Opin. Neurobiol. 10:88-94). Plexins are neuronal cell surface molecules that mediate cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Plexins have been shown to be expressed in the receptors and neurons of



particular sensory systems (Ohta, K. et al. (1995) Cell 14:1189-1199). There is evidence that suggests that some plexins function to control motor and CNS axon guidance in the developing nervous system. Plexins, which themselves contain complete semaphorin domains, may be both the ancestors of classical semaphorins and binding partners for semaphorins (Winberg, M.L. et al (1998) Cell 95:903-916).

Human pregnancy-specific beta 1-glycoprotein (PSG) is a family of closely related glycoproteins of molecular weights of 72 KDa, 64KDa, 62KDa, and 54KDa. Together with the carcinoembryonic antigen, they comprise a subfamily within the immunoglobulin superfamily (Plouzek, C.A. and Chou, J.Y. (1991) Endocrinology 129:950-958) Different subpopulations of PSG have been found to be produced by the trophoblasts of the human placenta, and the amnionic and chorionic membranes (Plouzek, C.A. et al. (1993) Placenta 14:277-285).

Autocrine motility factor (AMF) is one of the motility cytokines regulating tumor cell migration; therefore identification of the signaling pathway coupled with it has critical importance. Autocrine motility factor receptor (AMFR) expression has been found to be associated with tumor progression in thymoma (Ohta Y. et al. (2000) Int. J. Oncol. 17:259-264). AMFR is a cell surface glycoprotein of molecular weight 78KDa.

Hormones are signaling molecules that coordinately regulate basic physiological processes from embryogenesis throughout adulthood. These processes include metabolism, respiration, reproduction, excretion, fetal tissue differentiation and organogenesis, growth and development, homeostasis, and the stress response. Hormonal secretions and the nervous system are tightly integrated and interdependent. Hormones are secreted by endocrine glands, primarily the hypothalamus and pituitary, the thyroid and parathyroid, the pancreas, the adrenal glands, and the ovaries and testes.

The secretion of hormones into the circulation is tightly controlled. Hormones are often secreted in diurnal, pulsatile, and cyclic patterns. Hormone secretion is regulated by perturbations in blood biochemistry, by other upstream-acting hormones, by neural impulses, and by negative feedback loops. Blood hormone concentrations are constantly monitored and adjusted to maintain optimal, steady-state levels. Once secreted, hormones act only on those target cells that express specific receptors.

Most disorders of the endocrine system are caused by either hyposecretion or hypersecretion of hormones. Hyposecretion often occurs when a hormone's gland of origin is damaged or otherwise impaired. Hypersecretion often results from the proliferation of tumors derived from hormone-secreting cells. Inappropriate hormone levels may also be caused by defects in regulatory feedback loops or in the processing of hormone precursors. Endocrine malfunction may also occur when the target cell fails to respond to the hormone.

Hormones can be classified biochemically as polypeptides, steroids, eicosanoids, or amines. Polypeptide hormones, which include diverse hormones such as insulin and growth hormone, vary in size and function and are often synthesized as inactive precursors that are processed intracellularly into mature, active forms. Amine hormones, which include epinephrine and dopamine, are amino acid derivatives that function in neuroendocrine signaling. Steroid hormones, which include the cholesterol-derived hormones estrogen and testosterone, function in sexual development and reproduction. Eicosanoid hormones, which include prostaglandins and prostacyclins, are fatty acid derivatives that function in a variety of processes. Most polypeptide hormones and some amine hormones are soluble in the circulation where they are highly susceptible to proteolytic degradation within seconds after their secretion. Steroid hormones and eicosanoid hormones are insoluble and must be transported in the circulation by carrier proteins. The following discussion will focus primarily on polypeptide hormones.

Hormones secreted by the hypothalamus and pituitary gland play a critical role in endocrine function by coordinately regulating hormonal secretions from other endocrine glands in response to neural signals. Hypothalamic hormones include thyrotropin-releasing hormone, gonadotropin-releasing hormone, somatostatin, growth-hormone releasing factor, corticotropin-releasing hormone, substance P, dopamine, and prolactin-releasing hormone. These hormones directly regulate the secretion of hormones from the anterior lobe of the pituitary. Hormones secreted by the anterior pituitary include adrenocorticotrophic hormone (ACTH), melanocyte-stimulating hormone, somatotrophic hormones such as growth hormone and prolactin, glycoprotein hormones such as thyroid-stimulating hormone, luteinizing hormone (LH), and follicle-stimulating hormone (FSH),  $\beta$ -lipotropin, and  $\beta$ -endorphins. These hormones regulate hormonal secretions from the thyroid, pancreas, and adrenal glands, and act directly on the reproductive organs to stimulate ovulation and spermatogenesis. The posterior pituitary synthesizes and secretes antidiuretic hormone (ADH, vasopressin) and oxytocin.

Disorders of the hypothalamus and pituitary often result from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma. Such disorders have profound effects on the function of other endocrine glands. Disorders associated with hypopituitarism include hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism. Disorders associated with hyperpituitarism include acromegaly, gigantism, and syndrome of inappropriate ADH secretion (SIADH), often caused by benign adenomas.

Hormones secreted by the thyroid and parathyroid primarily control metabolic rates and the regulation of serum calcium levels, respectively. Thyroid hormones include calcitonin, somatostatin,

and thyroid hormone. The parathyroid secretes parathyroid hormone. Disorders associated with hypothyroidism include goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism. Disorders associated with hyperthyroidism include thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease. Disorders associated with hyperparathyroidism include Conn disease (chronic hypercalcemia) leading to bone resorption and parathyroid hyperplasia.

Hormones secreted by the pancreas regulate blood glucose levels by modulating the rates of carbohydrate, fat, and protein metabolism. Pancreatic hormones include insulin, glucagon, amylin,  $\gamma$ -aminobutyric acid, gastrin, somatostatin, and pancreatic polypeptide. The principal disorder associated with pancreatic dysfunction is diabetes mellitus caused by insufficient insulin activity. Diabetes mellitus is generally classified as either Type I (insulin-dependent, juvenile diabetes) or Type II (non-insulin-dependent, adult diabetes). The treatment of both forms by insulin replacement therapy is well known. Diabetes mellitus often leads to acute complications such as hypoglycemia (insulin shock), coma, diabetic ketoacidosis, lactic acidosis, and chronic complications leading to disorders of the eye, kidney, skin, bone, joint, cardiovascular system, nervous system, and to decreased resistance to infection.

The anatomy, physiology, and diseases related to hormonal function are reviewed in McCance, K. L. and Huether, S. E. (1994) Pathophysiology: The Biological Basis for Disease in Adults and Children, Mosby-Year Book, Inc., St. Louis, MO; Greenspan, F. S. and Baxter, J. D. (1994) Basic and Clinical Endocrinology, Appleton and Lange, East Norwalk, CT.

Growth factors are secreted proteins that mediate intercellular communication. Unlike hormones, which travel great distances via the circulatory system, most growth factors are primarily local mediators that act on neighboring cells. Most growth factors contain a hydrophobic N-terminal signal peptide sequence which directs the growth factor into the secretory pathway. Most growth factors also undergo post-translational modifications within the secretory pathway. These modifications can include proteolysis, glycosylation, phosphorylation, and intramolecular disulfide bond formation. Once secreted, growth factors bind to specific receptors on the surfaces of neighboring target cells, and the bound receptors trigger intracellular signal transduction pathways. These signal transduction pathways elicit specific cellular responses in the target cells. These responses can include the modulation of gene expression and the stimulation or inhibition of cell division, cell differentiation, and cell motility.

Growth factors fall into at least two broad and overlapping classes. The broadest class includes the large polypeptide growth factors, which are wide-ranging in their effects. These factors include epidermal growth factor (EGF), fibroblast growth factor (FGF), transforming growth factor- $\beta$

(TGF- $\beta$ ), insulin-like growth factor (IGF), nerve growth factor (NGF), and platelet-derived growth factor (PDGF), each defining a family of numerous related factors. The large polypeptide growth factors, with the exception of NGF, act as mitogens on diverse cell types to stimulate wound healing, bone synthesis and remodeling, extracellular matrix synthesis, and proliferation of epithelial, epidermal, and connective tissues. Members of the TGF- $\beta$ , EGF, and FGF families also function as inductive signals in the differentiation of embryonic tissue. NGF functions specifically as a neurotrophic factor, promoting neuronal growth and differentiation.

Another class of growth factors includes the hematopoietic growth factors, which are narrow in their target specificity. These factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. These factors include the colony-stimulating factors (G-CSF, M-CSF, GM-CSF, and CSF1-3), erythropoietin, and the cytokines. The cytokines are specialized hematopoietic factors secreted by cells of the immune system and are discussed in detail below.

Hormones travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical compositions and mechanisms of action, hormones can be grouped into two categories. One category includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines (epinephrine, norepinephrine) and histamine, and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotrophic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864.)

Pro-opiomelanocortin (POMC) is the precursor polypeptide of corticotropin (ACTH), a hormone synthesized by the anterior pituitary gland, which functions in the stimulation of the adrenal cortex. POMC is also the precursor polypeptide of the hormone beta-lipotropin (beta-LPH). Each hormone includes smaller peptides with distinct biological activities: alpha-melanotropin (alpha-MSH) and corticotropin-like intermediate lobe peptide (CLIP) are formed from ACTH; gamma-lipotropin (gamma-LPH) and beta-endorphin are peptide components of beta-LPH; while beta-MSH is contained within gamma-LPH. Adrenal insufficiency due to ACTH deficiency, resulting from a genetic mutation in exons 2 and 3 of POMC results in an endocrine disorder characterized by early-

onset obesity, adrenal insufficiency, and red hair pigmentation (Chretien, M. et al. (1979) Can. J. Biochem. 57:1111-1121; Krude, H. et al. (1998) Nat. Genet. 19:155-157; Online Mendelian Inheritance in Man (OMIM) 176830).

Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with membrane proteins for activity. Complex interactions among these factors and their receptors trigger intracellular signal transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth factors such as epidermal growth factor (EGF), fibroblast growth factor, transforming growth factor, insulin-like growth factor (IGF), and platelet-derived growth factor. EGF includes a 30-40 residue EGF repeat domain, composed of conserved cysteine and glycine residues, found in a variety of proteins involved in cell proliferation, including the leukocyte antigen CD97 and the Notch family proteins (Greener, M. (2000) Mol. Med. Today 6:139-140). IGF forms a heterotrimeric complex with IGF-binding-protein 3 and the acid-labile subunit (ALS). ALS is largely composed of 18-20 leucine-rich repeats of 24 amino acids (Leong, S.R. et al. (1992) Mol. Endocrinol. 6:870-876). The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs). Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin, which function as hormones to regulate cellular functions other than proliferation.

Growth and differentiation factors play critical roles in neoplastic transformation of cells in vitro and in tumor progression in vivo. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both in vivo and in vitro. Moreover, some growth factors and growth factor receptors are related both structurally and functionally to oncoproteins. In addition, growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)

In addition, some of the large polypeptide growth factors play crucial roles in the induction of the primordial germ layers in the developing embryo. This induction ultimately results in the formation of the embryonic mesoderm, ectoderm, and endoderm which in turn provide the framework for the entire adult body plan. Disruption of this inductive process would be catastrophic

to embryonic development. One such growth factor, wnt, is a secreted glycoprotein that has activity as both a short-range inducer and as a long-range morphogen (for a review, see Howes, R. and S. Bray (2000) *Current Biology* 10:R222-R226). Wnt signaling is implicated in diseases including cancer and Alzheimer's Disease (Bienz, M. and H. Clevers (2000) *Cell* 103:311-320; Polakis, P. (2000) *Genes Dev.* 14:1837-1851; De Ferrari, G.V. and N.C. Inestrosa (2000) *Brain Res. Brain. Res. Rev.* 33:1-12). Chordin is a developmental protein that binds to ventralizing TGF-beta-like bone morphogenetic proteins (BMPs) and sequesters them in latent complexes, causing dorsalization of tissue (Pappano, W. N. et al. (1998) *Genomics* 52:236-239). Other developmental proteins that regulate BMPs include noggin, cerberus, dan, and gremlin (Schmitt, J.M. et al. (1999) *J. Orthop. Res.* 17:269-278).

The Slit protein, first identified in *Drosophila*, is critical in central nervous system midline formation and potentially in nervous tissue histogenesis and axonal pathfinding. Itoh et al. ((1998) *Brain Res. Mol. Brain Res.* 62:175-186) have identified mammalian homologues of the slit gene (human Slit-1, Slit-2, Slit-3 and rat Slit-1). The encoded proteins are putative secreted proteins containing EGF-like motifs and leucine-rich repeats, both of which are conserved protein-protein interaction domains. Slit-1, -2, and -3 mRNAs are expressed in the brain, spinal cord, and thyroid, respectively (Itoh, A. et al., *supra*). The Slit family of proteins are indicated to be functional ligands of glypican-1 in nervous tissue and it is suggested that their interactions may be critical in certain stages during central nervous system histogenesis (Liang, Y. et al. (1999) *J. Biol. Chem.* 274:17885-17892).

Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C.R. et al. (1985) *Endocrine Physiology*, Oxford University Press, New York, NY, pp. 57-62.)

NP/VMs are involved in numerous neurological and cardiovascular disorders. For example, neuropeptide Y is involved in hypertension, congestive heart failure, affective disorders, and appetite regulation. Somatostatin inhibits secretion of growth hormone and prolactin in the anterior pituitary, as well as inhibiting secretion in intestine, pancreatic acinar cells, and pancreatic beta-cells. A reduction in somatostatin levels has been reported in Alzheimer's disease and Parkinson's disease.

Vasopressin acts in the kidney to increase water and sodium absorption, and in higher concentrations stimulates contraction of vascular smooth muscle, platelet activation, and glycogen breakdown in the liver. Vasopressin and its analogues are used clinically to treat diabetes insipidus. Endothelin and angiotensin are involved in hypertension, and drugs, such as captopril, which reduce plasma levels of angiotensin, are used to reduce blood pressure (Watson, S. and S. Arkinstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 194; 252; 284; 55; 111).

Neuropeptides have also been shown to have roles in nociception (pain). Vasoactive intestinal peptide appears to play an important role in chronic neuropathic pain. Nociceptin, an endogenous ligand for the opioid receptor-like 1 receptor, is thought to have a predominantly anti-nociceptive effect, and has been shown to have analgesic properties in different animal models of tonic or chronic pain (Dickinson, T. and Fleetwood-Walker, S.M. (1998) *Trends Pharmacol. Sci.* 19:346-348).

Cytokines comprise a family of signaling molecules that modulate the immune system and the inflammatory response. Cytokines are usually secreted by leukocytes, or white blood cells, in response to injury or infection. Cytokines function as growth and differentiation factors that act primarily on cells of the immune system such as B- and T-lymphocytes, monocytes, macrophages, and granulocytes. Like other signaling molecules, cytokines bind to specific plasma membrane receptors and trigger intracellular signal transduction pathways which alter gene expression patterns. There is considerable potential for the use of cytokines in the treatment of inflammation and immune system disorders.

Cytokine structure and function have been extensively characterized in vitro. Most cytokines are small polypeptides of about 30 kilodaltons or less. Over 50 cytokines have been identified from human and rodent sources. Examples of cytokine subfamilies include the interferons (IFN- $\alpha$ , - $\beta$ , and - $\gamma$ ), the interleukins (IL1-IL13), the tumor necrosis factors (TNF- $\alpha$  and - $\beta$ ), and the chemokines. Many cytokines have been produced using recombinant DNA techniques, and the activities of individual cytokines have been determined in vitro. These activities include regulation of leukocyte proliferation, differentiation, and motility.

The activity of an individual cytokine in vitro may not reflect the full scope of that cytokine's activity in vivo. Cytokines are not expressed individually in vivo but are instead expressed in combination with a multitude of other cytokines when the organism is challenged with a stimulus. Together, these cytokines collectively modulate the immune response in a manner appropriate for that particular stimulus. Therefore, the physiological activity of a cytokine is determined by the stimulus itself and by complex interactive networks among co-expressed cytokines which may demonstrate both synergistic and antagonistic relationships.

Chemokines comprise a cytokine subfamily with over 30 members. (Reviewed in Wells, T.

N. C. and Peitsch, M. C. (1997) J. Leukoc. Biol. 61:545-550.) Chemokines were initially identified as chemotactic proteins that recruit monocytes and macrophages to sites of inflammation. Recent evidence indicates that chemokines may also play key roles in hematopoiesis and HIV-1 infection. Chemokines are small proteins which range from about 6-15 kilodaltons in molecular weight.

- 5 Chemokines are further classified as C, CC, CXC, or CX<sub>3</sub>C based on the number and position of critical cysteine residues. The CC chemokines, for example, each contain a conserved motif consisting of two consecutive cysteines followed by two additional cysteines which occur downstream at 24- and 16-residue intervals, respectively (ExpASY PROSITE database, documents PS00472 and PDOC00434). The presence and spacing of these four cysteine residues are highly
- 10 conserved, whereas the intervening residues diverge significantly. However, a conserved tyrosine located about 15 residues downstream of the cysteine doublet seems to be important for chemotactic activity. Most of the human genes encoding CC chemokines are clustered on chromosome 17, although there are a few examples of CC chemokine genes that map elsewhere. Other chemokines include lymphotactin (C chemokine); macrophage chemotactic and activating factor (MCAF/MCP-1;
- 15 CC chemokine); platelet factor 4 and IL-8 (CXC chemokines); and fractalkine and neurotractin (CX<sub>3</sub>C chemokines). (Reviewed in Luster, A. D. (1998) N. Engl. J. Med. 338:436-445.)

- Other proteins that contain signal peptides include secreted proteins with enzymatic activity. Such activity includes, for example, oxidoreductase/dehydrogenase activity, transferase activity, hydrolase activity, lyase activity, isomerase activity, or ligase activity. For example, matrix
- 20 metalloproteinases are secreted hydrolytic enzymes that degrade the extracellular matrix and thus play an important role in tumor metastasis, tissue morphogenesis, and arthritis (Reponen, P. et al. (1995) Dev. Dyn. 202:388-396; Firestein, G.S. (1992) Curr. Opin. Rheumatol. 4:348-354; Ray, J.M. and Stetler-Stevenson, W.G. (1994) Eur. Respir. J. 7:2062-2072; and Mignatti, P. and Rifkin, D.B. (1993) Physiol. Rev. 73:161-195). The catalytic protein disulfide isomerase (PDI) is found in
- 25 membrane-bound eukaryotic compartments such as the endoplasmic reticulum (ER). It facilitates disulfide bond exchange as well as correct glycosylation. Edman et al. (1995; Nature 317:267-70) reported that rat PDI is useful for the *in vitro* production and folding of recombinant human proteins. Likewise, purified PDI is also commercially useful for the production and folding of recombinant, therapeutic human proteins such as tissue plasminogen activator (tPA). Ceruloplasmin is a serum
- 30 multicopper oxidase which plays a role in iron metabolism. Aceruloplasminemia is characterized by diabetes, retinal degeneration, and neurologic symptoms (for a review, see Gitlin, J.D. (1998) Pediatr. Res. 4:271-276). Additional examples are the acetyl-CoA synthetases which activate acetate for use in lipid synthesis or energy generation (Luong, A. et al. (2000) J. Biol. Chem. 275:26458-26466). The result of acetyl-CoA synthetase activity is the formation of acetyl-CoA from acetate and CoA.
- 35 Acetyl-CoA synthetases share a region of sequence similarity identified as the AMP-binding domain



signature. Acetyl-CoA synthetase has been shown to be associated with hypertension (Toh, H. (1991) Protein Seq. Data Anal. 4:111-117; and Iwai, N. et al. (1994) Hypertension 23:375-380).

A number of isomerases catalyze steps in protein folding, phototransduction, and various anabolic and catabolic pathways. One class of isomerases is known as peptidyl-prolyl *cis-trans* isomerases (PPIases). PPIases catalyze the *cis* to *trans* isomerization of certain proline imidic bonds in proteins. Two families of PPIases are the FK506 binding proteins (FKBPs), and cyclophilins (CyPs). FKBPs bind the potent immunosuppressants FK506 and rapamycin, thereby inhibiting signaling pathways in T-cells. Specifically, the PPIase activity of FKBPs is inhibited by binding of FK506 or rapamycin. There are five members of the FKBP family which are named according to their calculated molecular masses (FKBP12, FKBP13, FKBP25, FKBP52, and FKBP65), and localized to different regions of the cell where they associate with different protein complexes (Coss, M. et al. (1995) J. Biol. Chem. 270:29336-29341; Schreiber, S.L. (1991) Science 251:283-287).

The peptidyl-prolyl isomerase activity of CyP may be part of the signaling pathway that leads to T-cell activation. CyP isomerase activity is associated with protein folding and protein trafficking, and may also be involved in assembly/disassembly of protein complexes and regulation of protein activity. For example, in *Drosophila*, the CyP NinaA is required for correct localization of rhodopsins, while a mammalian CyP (Cyp40) is part of the Hsp90/Hsc70 complex that binds steroid receptors. The mammalian CypA has been shown to bind the gag protein from human immunodeficiency virus 1 (HIV-1), an interaction that can be inhibited by cyclosporin. Since cyclosporin has potent anti-HIV-1 activity, CypA may play an essential function in HIV-1 replication. Finally, Cyp40 has been shown to bind and inactivate the transcription factor c-Myb, an effect that is reversed by cyclosporin. This effect implicates CyPs in the regulation of transcription, transformation, and differentiation (Bergsma, D.J. et al (1991) J. Biol. Chem. 266:23204-23214; Hunter, T. (1998) Cell 92:141-143; and Leveron, J.D. and Ness, S.A. (1998) Mol. Cell. 1:203-211).

Gamma-carboxyglutamic acid (Gla) proteins rich in proline (PRGPs) are members of a family of vitamin K-dependent single-pass integral membrane proteins. These proteins are characterized by an extracellular amino terminal domain of approximately 45 amino acids rich in Gla. The intracellular carboxyl terminal region contains one or two copies of the sequence PPXY, a motif present in a variety of proteins involved in such diverse cellular functions as signal transduction, cell cycle progression, and protein turnover (Kulman, J.D. et al. (2001) Proc. Natl. Acad. Sci. USA 98:1370-1375). The process of post-translational modification of glutamic residues to form Gla is Vitamin K-dependent carboxylation. Proteins which contain Gla include plasma proteins involved in blood coagulation. These proteins are prothrombin, proteins C, S, and Z, and coagulation factors VII, IX, and X. Osteocalcin (bone-Gla protein, BGP) and matrix Gla-protein (MGP) also contain Gla (Friedman, P.A. and C.T. Przysiecki (1987) Int. J. Biochem. 19:1-7; C. Vermeer (1990) Biochem. J.

266:625-636).

### Immunoglobulins

Antigen recognition molecules are key players in the sophisticated and complex immune systems which all vertebrates have developed to provide protection from viral, bacterial, fungal, and parasitic infections. A key feature of the immune system is its ability to distinguish foreign molecules, or antigens, from "self" molecules. This ability is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as lymphocytes, granulocytes, and monocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily, members of which contain one or more repeats of a conserved structural domain. This Ig domain is comprised of antiparallel  $\beta$  sheets joined by a disulfide bond in an arrangement called the Ig fold. The criteria for a protein to be a member of the Ig superfamily is to have one or more Ig domains, which are regions of 70-110 amino acid residues in length homologous to either Ig variable-like (V) or Ig constant-like (C) domains. Members of the Ig superfamily include antibodies (Ab), T cell receptors (TCRs), class I and II major histocompatibility (MHC) proteins and immune cell-specific surface markers such as the "cluster of differentiation" or CD antigens, CD2, CD3, CD4, CD8, poly-Ig receptors, Fc receptors, neural cell-adhesion molecule (NCAM) and platelet-derived growth factor receptor (PDGFR).

Ig domains (V and C) are regions of conserved amino acid residues that give a polypeptide a globular tertiary structure called an immunoglobulin (or antibody) fold, which consists of two approximately parallel layers of  $\beta$ -sheets. Conserved cysteine residues form an intrachain disulfide-bonded loop, 55-75 amino acid residues in length, which connects the two layers of  $\beta$ -sheets. Each  $\beta$ -sheet has three or four anti-parallel  $\beta$ -strands of 5-10 amino acid residues. Hydrophobic and hydrophilic interactions of amino acid residues within the  $\beta$ -strands stabilize the Ig fold (hydrophobic on inward facing amino acid residues and hydrophilic on the amino acid residues in the outward facing portion of the strands). A V domain consists of a longer polypeptide than a C domain, with an additional pair of  $\beta$ -strands in the Ig fold.

A consistent feature of Ig superfamily genes is that each sequence of an Ig domain is encoded by a single exon. It is possible that the superfamily evolved from a gene coding for a single Ig domain involved in mediating cell-cell interactions. New members of the superfamily then arose by exon and gene duplications. Modern Ig superfamily proteins contain different numbers of V and/or C domains. Another evolutionary feature of this superfamily is the ability to undergo DNA rearrangements, a unique feature retained by the antigen receptor members of the family.

Many members of the Ig superfamily are integral plasma membrane proteins with extracellular Ig domains. The hydrophobic amino acid residues of their transmembrane domains and

their cytoplasmic tails are very diverse, with little or no homology among Ig family members or to known signal-transducing structures. There are exceptions to this general superfamily description. For example, the cytoplasmic tail of PDGFR has tyrosine kinase activity. In addition Thy-1 is a glycoprotein found on thymocytes and T cells. This protein has no cytoplasmic tail, but is instead  
5 attached to the plasma membrane by a covalent glycoposphatidylinositol linkage.

Another common feature of many Ig superfamily proteins is the interactions between Ig domains which are essential for the function of these molecules. Interactions between Ig domains of a multimeric protein can be either homophilic or heterophilic (i.e., between the same or different Ig domains). Antibodies are multimeric proteins which have both homophilic and heterophilic  
10 interactions between Ig domains. Pairing of constant regions of heavy chains forms the Fc region of an antibody and pairing of variable regions of light and heavy chains form the antigen binding site of an antibody. Heterophilic interactions also occur between Ig domains of different molecules. These interactions provide adhesion between cells for significant cell-cell interactions in the immune system and in the developing and mature nervous system. (Reviewed in Abbas, A.K. et al. (1991) Cellular and Molecular Immunology, W.B. Saunders Company, Philadelphia, PA, pp.142-145.)  
15

#### Antibodies

MHC proteins are cell surface markers that bind to and present foreign antigens to T cells. MHC molecules are classified as either class I or class II. Class I MHC molecules (MHC I) are expressed on the surface of almost all cells and are involved in the presentation of antigen to  
20 cytotoxic T cells. For example, a cell infected with virus will degrade intracellular viral proteins and express the protein fragments bound to MHC I molecules on the cell surface. The MHC I/antigen complex is recognized by cytotoxic T-cells which destroy the infected cell and the virus within. Class II MHC molecules are expressed primarily on specialized antigen-presenting cells of the immune system, such as B-cells and macrophages. These cells ingest foreign proteins from the  
25 extracellular fluid and express MHC II/antigen complex on the cell surface. This complex activates helper T-cells, which then secrete cytokines and other factors that stimulate the immune response. MHC molecules also play an important role in organ rejection following transplantation. Rejection occurs when the recipient's T-cells respond to foreign MHC molecules on the transplanted organ in the same way as to self MHC molecules bound to foreign antigen. (Reviewed in Alberts, B. et al.  
30 (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, pp. 1229-1246.)

Antibodies are multimeric members of the Ig superfamily which are either expressed on the surface of B-cells or secreted by B-cells into the circulation. Antibodies bind and neutralize foreign antigens in the blood and other extracellular fluids. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-  
35 chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to

antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the  $\alpha$ ,  $\delta$ ,  $\epsilon$ ,  $\gamma$ , and  $\mu$  H-chain types. There are two types of L-chains,  $\kappa$  and  $\lambda$ , either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H- and L-chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site. (Reviewed in Alberts, B. et al. *supra*, pp. 1206-1213 and 1216-1217.)

Both H-chains and L-chains contain the repeated Ig domains of members of the Ig superfamily. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region.

The immune system is capable of recognizing and responding to any foreign molecule that enters the body. Therefore, the immune system must be armed with a full repertoire of antibodies against all potential antigens. Such antibody diversity is generated by somatic rearrangement of gene segments encoding variable and constant regions. These gene segments are joined together by site-specific recombination which occurs between highly conserved DNA sequences that flank each gene segment. Because there are hundreds of different gene segments, millions of unique genes can be generated combinatorially. In addition, imprecise joining of these segments and an unusually high rate of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

The discovery of new secreted proteins, and the polynucleotides encoding them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

## SUMMARY OF THE INVENTION

The invention features purified polypeptides, secreted proteins, referred to collectively as "SECP" and individually as "SECP-1," "SECP-2," "SECP-3," "SECP-4," "SECP-5," "SECP-6," "SECP-7," "SECP-8," "SECP-9," "SECP-10," "SECP-11," "SECP-12," "SECP-13," "SECP-14," "SECP-15," "SECP-16," "SECP-17," "SECP-18," "SECP-19," "SECP-20," "SECP-21," "SECP-22," "SECP-23," "SECP-24," "SECP-25," "SECP-26," "SECP-27," "SECP-28," "SECP-29," "SECP-30," "SECP-31," "SECP-32," "SECP-33," "SECP-34," "SECP-35," "SECP-36," "SECP-37," "SECP-38," "SECP-39," "SECP-40," "SECP-41," "SECP-42," "SECP-43," "SECP-44," "SECP-45," "SECP-46," "SECP-47," "SECP-48," "SECP-49," "SECP-50," "SECP-51," "SECP-52," "SECP-53," and "SECP-54." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-54.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-54. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:55-108.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the

invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and

which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous  
5 nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at  
10 least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target  
15 polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected  
20 from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-  
25 54. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino  
30 acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group  
35 consisting of SEQ ID NO:1-54. The method comprises a) exposing a sample comprising the

polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group



consisting of SEQ ID NO:1-54. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

## BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

## DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing

the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

## DEFINITIONS

5           “SECP” refers to the amino acid sequences of substantially purified SECP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

          The term “agonist” refers to a molecule which intensifies or mimics the biological activity of SECP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other  
10   compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

          An “allelic variant” is an alternative form of the gene encoding SECP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or  
15   many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

          “Altered” nucleic acid sequences encoding SECP include those sequences with deletions,  
20   insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as SECP or a polypeptide with at least one functional characteristic of SECP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding SECP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding  
25   SECP. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent SECP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of SECP is retained. For example,  
30   negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

35           The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide,

polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

5 "Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of SECP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small  
10 molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant.  
15 Antibodies that bind SECP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin,  
20 and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures  
25 on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an in vitro evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No.  
30 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH<sub>2</sub>), which may improve a desired property, e.g.,  
35 resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules,

e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

The term "intramer" refers to an aptamer which is expressed in vivo. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl Acad. Sci. USA 96:3606-3610).

The term "spiegelmer" refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic SECP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding SECP or fragments of SECP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be

deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide.

Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A "fragment" is a unique portion of SECP or the polynucleotide encoding SECP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:55-108 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:55-108, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:55-108 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:55-108 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:55-108 and the region of SEQ ID NO:55-108 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-54 is encoded by a fragment of SEQ ID NO:55-108. A

fragment of SEQ ID NO:1-54 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-54. For example, a fragment of SEQ ID NO:1-54 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-54. The precise length of a fragment of SEQ ID NO:1-54 and the region of SEQ ID NO:1-54 to which the  
5 fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

10 "Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps  
15 in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of  
20 molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent  
25 similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at  
30 <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>.  
35 The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST



programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

5 *Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

*Expect: 10*

10 *Word Size: 11*

*Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at  
15 least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode  
20 similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a  
25 standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default  
30 parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by  
35 CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

5        *Matrix: BLOSUM62*  
      *Open Gap: 11 and Extension Gap: 1 penalties*  
      *Gap x drop-off: 50*  
      *Expect: 10*  
      *Word Size: 3*  
10       *Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least  
15 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for  
20 chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a  
25 complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific  
30 binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about  
35 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g.,  $C_0t$  or  $R_0t$  analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of SECP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of SECP which is useful in any of the antibody production methods disclosed herein or known in the

art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

5 The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of SECP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of SECP.

10 The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably  
15 linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of  
20 amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an SECP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in  
25 the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of SECP.

"Probe" refers to nucleic acid sequences encoding SECP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule.

30 Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

35 Probes and primers as used in the present invention typically comprise at least 15 contiguous

nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing SECP, nucleic acids encoding SECP, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

5 A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells,  
10 trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods  
15 well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of  
20 replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the  
25 art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria,  
30 cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

35 A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having

at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

## THE INVENTION

The invention is based on the discovery of new human secreted proteins (SECP), the polynucleotides encoding SECP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is



denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs, including the locations of signal peptides (as indicated by "Signal Peptide" and/or "signal\_cleavage"). Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are secreted proteins. For example, SEQ ID NO:2 is 99% identical to a novel human AMP-binding enzyme similar to acetyl-coenzyme A synthetase (acetate-coA ligase) (GenBank ID g6996429) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is  $5.8e-262$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains an AMP-binding domain signature as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:2 is an AMP-binding enzyme (note that "AMP-binding domains" are shared regions of sequence similarity within a number of prokaryotic and eukaryotic enzymes which most likely act via an ATP-dependent covalent binding of AMP to their substrate, PROSITE:PDOC00427).

As a further example, SEQ ID NO:3 is 33% identical from residues E44 to L530 to bovine PDI (protein disulfide isomerase) (GenBank ID g163497) as determined by the Basic Local

Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is  $1.1e-70$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:3 also contains a thioredoxin domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:3 is a protein disulfide isomerase.

As a further example, SEQ ID NO:4 is 56% identical to human preceruloplasmin (GenBank ID g180256) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:4 contains a signal peptide and a multicopper oxidase active site domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) The presence of this domain is confirmed by BLIMPS, MOTIFS, and PROFILESCAN analyses, providing further corroborative evidence that SEQ ID NO:4 is a secreted multicopper oxidase.

In another example, SEQ ID NO:16 is 79% identical to human growth hormone hGH-V2 (GenBank ID g183178) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is  $5.6e-106$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:16 also contains a signal peptide and a somatotropin hormone family signature as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) The presence of these motifs is confirmed by BLIMPS, MOTIFS, SPSCAN, and PROFILESCAN analyses, providing further corroborative evidence that SEQ ID NO:16 is a secreted hormone.

As a further example, SEQ ID NO:27 is 49% identical to mouse Fca/m receptor (GenBank ID g11071950) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is  $2.2e-115$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:27 also contains an immunoglobulin domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from additional BLAST analyses provide further corroborative evidence that SEQ ID NO:27 is an immunoglobulin domain-containing receptor.

In another example, SEQ ID NO:41 is 99% identical to human chordin (GenBank ID g3822218) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:41 also contains a von Willebrand factor growth

regulator domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from MOTIFS analyses provide further corroborative evidence that SEQ ID NO:41 is a growth regulation molecule.

5 SEQ ID NO:50 contains a signal peptide as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 2.) The presence of the signal peptide is confirmed by data from SPSCAN. SEQ ID NO:1, SEQ ID NO:5-15, SEQ ID NO:17-26, SEQ ID NO:28-40, SEQ ID NO:42-49 and SEQ ID NO:51-54, which were analyzed and annotated in a similar manner, all contain signal  
10 peptides as determined by SPSCAN or HMMER analysis. The algorithms and parameters for the analysis of SEQ ID NO:1-54 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence  
15 identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide sequences of the invention, and of fragments of the polynucleotide sequences which are useful, for  
20 example, in hybridization or amplification technologies that identify SEQ ID NO:55-108 or that distinguish between SEQ ID NO:55-108 and related polynucleotide sequences.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank  
25 cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences  
30 including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as  
FL\_XXXXXX\_N<sub>1</sub>\_N<sub>2</sub>\_YYYYY\_N<sub>3</sub>\_N<sub>4</sub> represents a "stitched" sequence in which XXXXXX is the  
35 identification number of the cluster of sequences to which the algorithm was applied, and YYYYY is

the number of the prediction generated by the algorithm, and  $N_{1,2,3,\dots}$ , if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as

- 5 FLXXXXXX\_gAAAAA\_gBBBBB\_1\_N is a "stretched" sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and  $N$  referring to specific exons (See Example V). In instances where a RefSeq  
10 sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (i.e., gBBBBB).

- Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis  
15 methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
20 FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

- In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA  
25 identification numbers are not shown.

- Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors  
30 which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses SECP variants. A preferred SECP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid

sequence identity to the SECP amino acid sequence, and which contains at least one functional or structural characteristic of SECP.

The invention also encompasses polynucleotides which encode SECP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:55-108, which encodes SECP. The polynucleotide sequences of SEQ ID NO:55-108, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding SECP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding SECP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:55-108 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:55-108. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SECP.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide sequence encoding SECP. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding SECP, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to the polynucleotide sequence encoding SECP over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding SECP. For example, a polynucleotide comprising a sequence of SEQ ID NO:108 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:94. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SECP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding SECP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These

combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring SECP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode SECP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring SECP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding SECP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SECP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode SECP and SECP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding SECP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:55-108 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences

are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding SECP may be extended utilizing a partial nucleotide  
5 sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown  
10 sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme  
15 digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries  
20 and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

25 When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

30 Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate  
35 software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire

process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof  
5 which encode SECP may be cloned in recombinant DNA molecules that direct expression of SECP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express SECP.

The nucleotide sequences of the present invention can be engineered using methods generally  
10 known in the art in order to alter SECP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction  
15 sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or  
20 improve the biological properties of SECP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of  
25 DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby  
30 maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding SECP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)  
35 Alternatively, SECP itself or a fragment thereof may be synthesized using chemical methods. For



example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of SECP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.)

The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active SECP, the nucleotide sequences encoding SECP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding SECP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding SECP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding SECP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding SECP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding SECP. These include, but are not limited to, microorganisms such as bacteria transformed

with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Bulter, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding SECP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding SECP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding SECP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of SECP are needed, e.g. for the production of antibodies, vectors which direct high level expression of SECP may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of SECP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of SECP. Transcription of sequences encoding SECP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding SECP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses SECP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of SECP in cell lines is preferred. For example, sequences encoding SECP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic,

or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding SECP is inserted within a marker gene sequence, transformed cells containing sequences encoding SECP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding SECP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding SECP and that express SECP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of SECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SECP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled

hybridization or PCR probes for detecting sequences related to polynucleotides encoding SECP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding SECP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding SECP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode SECP may be designed to contain signal sequences which direct secretion of SECP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding SECP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric SECP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of SECP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and

metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SECP encoding sequence and the heterologous protein sequence, so that SECP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled SECP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, <sup>35</sup>S-methionine.

SECP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to SECP. At least one and up to a plurality of test compounds may be screened for specific binding to SECP. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of SECP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which SECP binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express SECP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing SECP or cell membrane fractions which contain SECP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either SECP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with SECP, either in solution or affixed to a solid support, and detecting the binding of SECP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a

solid support.

SECP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of SECP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for SECP activity, wherein SECP is combined with at least one test compound, and the activity of SECP in the presence of a test compound is compared with the activity of SECP in the absence of the test compound. A change in the activity of SECP in the presence of the test compound is indicative of a compound that modulates the activity of SECP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising SECP under conditions suitable for SECP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of SECP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding SECP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding SECP may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding SECP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a

region of a polynucleotide encoding SECP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease.

- 5 Alternatively, a mammal inbred to overexpress SECP, e.g., by secreting SECP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

## THERAPEUTICS

- Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of SECP and secreted proteins. In addition, the expression of SECP is closely associated with breast, reproductive, digestive, urinary, fibroblastic, diseased, tumorous, testicular, 10 pituitary, adenoid, lymph node, monocyte, ileum, coronary artery endothelium, uterine endometrial and brain tissues. Examples can also be found in Table 6. Therefore, SECP appears to play a role in cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders. In the treatment of disorders associated with increased SECP expression or activity, it is 15 desirable to decrease the expression or activity of SECP. In the treatment of disorders associated with decreased SECP expression or activity, it is desirable to increase the expression or activity of SECP.

- Therefore, in one embodiment, SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or 20 activity of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of 25 the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, 30 autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, 35 hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or



pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, 5 fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus 10 erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; a neurological 15 disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, 20 suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central 25 nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), 30 akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, 35 and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary

mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss.

5        In another embodiment, a vector capable of expressing SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those described above.

10        In a further embodiment, a composition comprising a substantially purified SECP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those provided above.

15        In still another embodiment, an agonist which modulates the activity of SECP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those listed above.

20        In a further embodiment, an antagonist of SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders described above. In one aspect, an antibody which specifically binds SECP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express SECP.

25        In an additional embodiment, a vector expressing the complement of the polynucleotide encoding SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP including, but not limited to, those described above.

30        In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

35        An antagonist of SECP may be produced using methods which are generally known in the art. In particular, purified SECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SECP. Antibodies to SECP may also be generated using methods that are well known in the art. Such antibodies may include, but are not

limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with SECP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to SECP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of SECP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to SECP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SECP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for SECP may also be generated. For example, such fragments include, but are not limited to,  $F(ab')_2$  fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the  $F(ab')_2$  fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between SECP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering SECP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for SECP. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of SECP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple SECP epitopes, represents the average affinity, or avidity, of the antibodies for SECP. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular SECP epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the SECP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of SECP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of SECP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding SECP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding SECP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding SECP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding SECP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the

case where a genetic deficiency in SECP expression or regulation causes disease, the expression of SECP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SECP are treated by constructing mammalian expression vectors encoding SECP and introducing these vectors by mechanical means into SECP-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of SECP include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). SECP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or  $\beta$ -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding SECP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to SECP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding SECP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive

element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4<sup>+</sup> T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SECP to cells which have one or more genetic abnormalities with respect to the expression of SECP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding SECP to target cells which have one or more genetic abnormalities with respect to the expression of SECP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing SECP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S.

Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding SECP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for SECP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of SECP-coding RNAs and the synthesis of high levels of SECP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of SECP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using



triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

5 Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding SECP.

10 Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of  
15 candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis.

20 Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding SECP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs... that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

25 RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine,  
30 queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding SECP. Compounds  
35 which may be effective in altering expression of a specific polynucleotide may include, but are not

limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased SECP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding SECP may be therapeutically useful, and in the treatment of disorders associated with decreased SECP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding SECP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding SECP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding SECP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding SECP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable

for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat.

5 Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

10 An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of SECP, antibodies to SECP, and mimetics, agonists, antagonists, or inhibitors of SECP.

15 The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form.  
20 These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton,  
25 J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

30 Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising SECP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, SECP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to  
35 transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et

al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration  
5 range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example SECP or fragments thereof, antibodies of SECP, and agonists, antagonists or inhibitors of SECP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by  
10 standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the  $ED_{50}$  (the dose therapeutically effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the  $LD_{50}/ED_{50}$  ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are  
15 used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the  
20 subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week,  
25 or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1  $\mu\text{g}$  to 100,000  $\mu\text{g}$ , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their  
30 inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

## DIAGNOSTICS

In another embodiment, antibodies which specifically bind SECP may be used for the diagnosis of disorders characterized by expression of SECP, or in assays to monitor patients being  
35 treated with SECP or agonists, antagonists, or inhibitors of SECP. Antibodies useful for diagnostic

purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for SECP include methods which utilize the antibody and a label to detect SECP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring SECP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of SECP expression. Normal or standard values for SECP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to SECP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of SECP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding SECP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of SECP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of SECP, and to monitor regulation of SECP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding SECP or closely related molecules may be used to identify nucleic acid sequences which encode SECP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding SECP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the SECP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:55-108 or from genomic sequences including promoters, enhancers, and introns of the SECP gene.

Means for producing specific hybridization probes for DNAs encoding SECP include the cloning of polynucleotide sequences encoding SECP or SECP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a

variety of reporter groups, for example, by radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding SECP may be used for the diagnosis of disorders associated with expression of SECP. Examples of such disorders include, but are not limited to, a

5 cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall

10 bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune

15 polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis,

20 myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a

25 cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy,

30 myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; a neurological disorder such as epilepsy, ischemic cerebrovascular

35 disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease,

dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and  
5 radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal  
10 disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive  
15 dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-  
20 Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss. The polynucleotide sequences encoding SECP may be used in Southern or northern  
25 analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding SECP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide  
30 sequences encoding SECP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding SECP in the  
35 sample indicates the presence of the associated disorder. Such assays may also be used to evaluate

the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of SECP, a normal or standard profile for expression is established. This may be accomplished by combining  
5 body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding SECP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from  
10 samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from  
15 successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance  
20 of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding SECP may involve the use of PCR. These oligomers may be chemically synthesized, generated  
25 enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding SECP, or a fragment of a polynucleotide complementary to the polynucleotide encoding SECP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

30 In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding SECP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP,  
35 oligonucleotide primers derived from the polynucleotide sequences encoding SECP are used to



amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of SECP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, SECP, fragments of SECP, or antibodies specific for SECP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to

generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the

treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples  
5 are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles,  
10 are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by  
15 sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from  
20 biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the  
25 polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for SECP to quantify the levels of SECP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting  
30 the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoz, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

35 Toxicant signatures at the proteome level are also useful for toxicological screening, and

should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding SECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a

chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding SECP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, SECP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between SECP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with SECP, or fragments thereof, and washed. Bound SECP is then detected by methods well known in the art. Purified SECP can

also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

5 In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding SECP specifically compete with a test compound for binding SECP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with SECP.

10 In additional embodiments, the nucleotide sequences which encode SECP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure  
15 in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/262,932, U.S. Ser. No. 60/265,926, U.S. Ser. No. 60/255,639, U.S. Ser. No. 60/257,852, U.S. Ser. No. 60/260,105, U.S. Ser. No. 60/263,090 and U.S. Ser. No. 60/263,096 are expressly incorporated by reference herein.

20

## EXAMPLES

### I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of  
25 denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

30 Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA  
35 purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectromAX DH10B from Life Technologies.

## II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

## III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the

MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from Homo sapiens, Rattus norvegicus, Mus musculus, Caenorhabditis elegans, Saccharomyces cerevisiae, Schizosaccharomyces pombe, and Candida albicans (Incyte Genomics, Palo Alto CA); and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence



alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:55-108. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

#### **IV. Identification and Editing of Coding Sequences from Genomic DNA**

Putative secreted proteins were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode secreted proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted proteins. Potential secreted proteins were also identified by homology to Incyte cDNA sequences that had been annotated as secreted proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly

process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

#### V. Assembly of Genomic Sequence Data with cDNA Sequence Data

##### "Stitched" Sequences

5 Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpi public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

##### "Stretched" Sequences

25 Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The

35

resultant stretched sequences were examined to determine whether it contained a complete gene.

#### VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:55-108 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:55-108 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:58 was mapped to chromosome 3 within the interval from 160.0 to 187.1 centiMorgans. SEQ ID NO:59 was mapped to chromosome 15 within the interval from 59.3 centiMorgans to the q-terminus. SEQ ID NO:60 was mapped to chromosome 15 within the interval from 39.5 to 59.3 centiMorgans. SEQ ID NO:61 was mapped to chromosome 3 within the interval from 67.9 to 77.4 centiMorgans. SEQ ID NO:62 was mapped to chromosome 16 at 473.44 centiMorgans. SEQ ID NO:63 was mapped to chromosome 9 within the interval from 75.8 to 136.7 centiMorgans. SEQ ID NO:64 was mapped to chromosome 19. SEQ ID NO:65 was mapped to chromosome 1 within the interval from 196.5 to 205.1 centiMorgans. SEQ ID NO:66 was mapped to chromosome 5 within the interval from 138.7 to 141.4 centiMorgans. SEQ ID NO:67 was mapped to chromosome 2 within the interval from 223.1 to 231.8 centiMorgans. SEQ ID NO:68 was mapped to chromosome 2 within the interval from 223.1 to 231.8 centiMorgans. SEQ ID NO:69 was mapped to chromosome 17 within the interval from 62.2 centiMorgans to the q-terminus. SEQ ID NO:75 was mapped to chromosome 15 within the interval from 59.3 centiMorgans to the q terminus. SEQ ID NO:76 was mapped to chromosome 13 within the interval from the p-terminus to 36.6 centiMorgans.

SEQ ID NO:77 was mapped to the short arm of chromosome 8 within the cytogenetic band 23.3.  
SEQ ID NO:78 was mapped to chromosome 11 within the interval from 102.6 to 131.7 centiMorgans.  
SEQ ID NO:79 was mapped to chromosome 3 within the interval from 49.5 to 64.4 centiMorgans.  
SEQ ID NO:80 was mapped to chromosome 5 within the interval from 104.5 to 121.4 centiMorgans.

## 5 VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel (1995) supra, ch. 4 and 16.)

10 Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

15

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum \{length(Seq. 1), length(Seq. 2)\}}}$$

The product score takes into account both the degree of similarity between two sequences and the  
20 length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by  
25 gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the  
30 other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding SECP are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA  
35 sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is

classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding SECP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

### VIII. Extension of SECP Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar,

Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

#### **IX. Labeling and Use of Individual Hybridization Probes**

Hybridization probes derived from SEQ ID NO:55-108 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a

SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing  $10^7$  counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

5 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and  
10 compared.

#### X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the  
15 aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to  
20 those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be  
25 selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser  
30 desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

#### Tissue or Cell Sample Preparation

35 Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and

poly(A)<sup>+</sup> RNA is purified using the oligo-(dT) cellulose method. Each poly(A)<sup>+</sup> RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ $\mu$ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ $\mu$ l RNase inhibitor, 500  $\mu$ M dATP, 500  $\mu$ M dGTP, 500  $\mu$ M dTTP, 40  $\mu$ M dCTP, 40  $\mu$ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)<sup>+</sup> RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)<sup>+</sup> RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14  $\mu$ l 5X SSC/0.2% SDS.

#### 15 Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5  $\mu$ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1  $\mu$ l of the array element DNA, at an average concentration of 100 ng/ $\mu$ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in



0.2% SDS and distilled water as before.

### Hybridization

Hybridization reactions contain 9  $\mu$ l of sample mixture consisting of 0.2  $\mu$ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140  $\mu$ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

### Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital

(A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

#### XI. Complementary Polynucleotides

Sequences complementary to the SECP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring SECP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of SECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the SECP-encoding transcript.

#### XII. Expression of SECP

Expression and purification of SECP is achieved using bacterial or virus-based expression systems. For expression of SECP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express SECP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of SECP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding SECP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K.

et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, SECP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from SECP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified SECP obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII, where applicable.

### XIII. Functional Assays

SECP function is assessed by expressing the sequences encoding SECP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10  $\mu$ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu$ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are

discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of SECP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding SECP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding SECP and other genes of interest can be analyzed by northern analysis or microarray techniques.

#### 10 XIV. Production of SECP Specific Antibodies

SECP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the SECP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-SECP activity by, for example, binding the peptide or SECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

#### XV. Purification of Naturally Occurring SECP Using Specific Antibodies

Naturally occurring or recombinant SECP is substantially purified by immunoaffinity chromatography using antibodies specific for SECP. An immunoaffinity column is constructed by covalently coupling anti-SECP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing SECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SECP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt

antibody/SECP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and SECP is collected.

#### **XVI. Identification of Molecules Which Interact with SECP**

SECP, or biologically active fragments thereof, are labeled with  $^{125}\text{I}$  Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled SECP, washed, and any wells with labeled SECP complex are assayed. Data obtained using different concentrations of SECP are used to calculate values for the number, affinity, and association of SECP with the candidate molecules.

Alternatively, molecules interacting with SECP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

SECP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

#### **XVII. Demonstration of SECP Activity**

An assay for growth stimulating or inhibiting activity of SECP measures the amount of DNA synthesis in Swiss mouse 3T3 cells (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY). In this assay, varying amounts of SECP are added to quiescent 3T3 cultured cells in the presence of  $[^3\text{H}]$ thymidine, a radioactive DNA precursor. SECP for this assay can be obtained by recombinant means or from biochemical preparations. Incorporation of  $[^3\text{H}]$ thymidine into acid-precipitable DNA is measured over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold SECP concentration range is indicative of growth modulating activity. One unit of activity per milliliter is defined as the concentration of SECP producing a 50% response level, where 100% represents maximal incorporation of  $[^3\text{H}]$ thymidine into acid-precipitable DNA.

Alternatively, an assay for SECP activity measures the stimulation or inhibition of neurotransmission in cultured cells. Cultured CHO fibroblasts are exposed to SECP. Following endocytic uptake of SECP, the cells are washed with fresh culture medium, and a whole cell voltage-clamped *Xenopus* myocyte is manipulated into contact with one of the fibroblasts in SECP-free medium. Membrane currents are recorded from the myocyte. Increased or decreased current relative to control values are indicative of neuromodulatory effects of SECP (Morimoto, T. et al. (1995) *Neuron* 15:689-696).

Alternatively, an assay for SECP activity measures the amount of SECP in secretory, membrane-bound organelles. Transfected cells as described above are harvested and lysed. The lysate is fractionated using methods known to those of skill in the art, for example, sucrose gradient ultracentrifugation. Such methods allow the isolation of subcellular components such as the Golgi apparatus, ER, small membrane-bound vesicles, and other secretory organelles.

Immunoprecipitations from fractionated and total cell lysates are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The concentration of SECP in secretory organelles relative to SECP in total cell lysate is proportional to the amount of SECP in transit through the secretory pathway.

Alternatively, AMP binding activity is measured by combining SECP with  $^{32}\text{P}$ -labeled AMP. The reaction is incubated at  $37^\circ\text{C}$  and terminated by addition of trichloroacetic acid. The acid extract is neutralized and subjected to gel electrophoresis to remove unbound label. The radioactivity retained in the gel is proportional to SECP activity.

Alternatively, the activity of purified SECP can be tested by introducing the molecule into an *in vitro* production system for tissue plasminogen activator (tPA). Any statistically significant improvement of correctly folded tPA in the presence as compared to the absence of SECP would indicate that SECP is active and functioning correctly.

Alternatively, SECP activity may be measured by the enzymatic activity they possess. For SEQ ID NO:4, for example, SECP activity is measured as ferroxidase activity at pH 6 in 0.3 M acetate buffer. The appearance of ferric ions is monitored at 315 nm (Bonomi, F. et al. (1996) J. Biol. Inorg. Chem. 1:67-72). For SEQ ID NO:6, for example, SECP activity is measured by the phosphorylation of galactose. SECP is incubated for 5 minutes in a 100  $\mu\text{l}$  reaction containing 200  $\mu\text{M}$   $^3\text{H}$ -galactose (30,000 cpm), 5 mM ATP, 5 mM  $\text{MgCl}_2$ , 5 mM NaF, 100 mM Tris-HCl buffer, pH 8.5. The reaction is stopped by heating at  $100^\circ\text{C}$  for 1 min, and the incubation mixture applied to a DE52 column. The column is washed with at least 5 column volumes of 10 mM  $(\text{NH}_4)\text{HCO}_3$  to remove unbound material. Galactose-P is eluted with 500 mM  $(\text{NH}_4)\text{HCO}_3$  and assayed for radioactive content by scintillation counting (Pastuszak, I. et al. (1996) J. Biol. Chem. 271:23653-23656). For SEQ ID NO:9, for example, SECP activity is measured by the amount of cobalamin bound using the isotope dilution method of Nexø, and Gimsing, employing human IF as the binding protein (1981, Scand. J. Clin. Lab. Invest. 41:465-468). For SEQ ID NO:10, for example, SECP activity is measured by the hydrolysis of appropriate synthetic peptide substrates conjugated with various chromogenic molecules in which the degree of hydrolysis is quantified by spectrophotometric (or fluorometric) absorption of the released chromophore (Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York, NY, pp.25-55). Peptide substrates are designed according to the category of protease activity as

endopeptidase (serine, cysteine, aspartic proteases, or metalloproteases), aminopeptidase (leucine aminopeptidase), or carboxypeptidase (carboxypeptidases A and B, procollagen C-proteinase). Commonly used chromogens are 2-naphthylamine, 4-nitroaniline, and furylacrylic acid. Assays are performed at ambient temperature and contain an aliquot of the enzyme and the appropriate substrate in a suitable buffer. Reactions are carried out in an optical cuvette, and the increase/decrease in absorbance of the chromogen released during hydrolysis of the peptide substrate is measured. The change in absorbance is proportional to the enzyme activity in the assay.

Alternatively, SECP activity can be measured as enzyme activity. For SEQ ID NO:20, for example, activity is proportional to the hydrolysis of glucosamine-6-sulfate by SECP which can be measured by the method of Robertson et al. (1992, Biochem. J. 288:539-544).

In another alternative, SECP can be assayed by its interaction with the insulin-like growth factor complex. For SEQ ID NO:17, for example, <sup>125</sup>I-labeled SECP is incubated for 2 h with 10 ng of IGF-I or-II and a range from 0 to 10 ng of IGFBP-3 in 50 mM sodium phosphate buffer, pH 6.5, at 22 °C (final volume 0.3 ml). SECP complexed to IGFBP-3 is precipitated using IGFBP-3 antiserum and radioactivity in each tube measured (Janosi, J.B.M. et al. (1999) J. Biol. Chem. 274:5292-5298).

#### XVIII. Demonstration of Immunoglobulin Activity

An assay for SECP activity measures the ability of SECP to recognize and precipitate antigens from serum. This activity can be measured by the quantitative precipitin reaction. (Golub, E.S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pages 113-115.)

SECP is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled SECP. SECP-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable SECP-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable SECP-antigen complex is plotted against the serum concentration. For various serum concentrations, a characteristic precipitin curve is obtained, in which the amount of precipitable SECP-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable SECP-antigen complex is a measure of SECP activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

Alternatively, an assay for SECP activity measures the expression of SECP on the cell surface. cDNA encoding SECP is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of SECP expressed on the cell surface.

Alternatively, an assay for SECP activity measures the amount of cell aggregation induced by overexpression of SECP. In this assay, cultured cells such as NIH3T3 are transfected with cDNA encoding SECP contained within a suitable mammalian expression vector under control of a strong promoter. Cotransfection with cDNA encoding a fluorescent marker protein, such as Green  
5 Fluorescent Protein (CLONTECH), is useful for identifying stable transfectants. The amount of cell agglutination, or clumping, associated with transfected cells is compared with that associated with untransfected cells. The amount of cell agglutination is a direct measure of SECP activity.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the  
10 invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.



Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
95765	1	095765CD1	55	095765CB1
6399886	2	6399886CD1	56	6399886CB1
6024420	3	6024420CD1	57	6024420CB1
7481067	4	7481067CD1	58	7481067CB1
3378720	5	3378720CD1	59	3378720CB1
938824	6	938824CD1	60	938824CB1
1683721	7	1683721CD1	61	1683721CB1
1694122	8	1694122CD1	62	1694122CB1
1970615	9	1970615CD1	63	1970615CB1
2314152	10	2314152CD1	64	2314152CB1
2886225	11	2886225CD1	65	2886225CB1
6144418	12	6144418CD1	66	6144418CB1
6834184	13	6834184CD1	67	6834184CB1
6951005	14	6951005CD1	68	6951005CB1
7250331	15	7250331CD1	69	7250331CB1
1758413	16	1758413CD1	70	1758413CB1
7011042	17	7011042CD1	71	7011042CB1
7427362	18	7427362CD1	72	7427362CB1
7485304	19	7485304CD1	73	7485304CB1
1422394	20	1422394CD1	74	1422394CB1
1336022	21	1336022CD1	75	1336022CB1
7473674	22	7473674CD1	76	7473674CB1
7475846	23	7475846CD1	77	7475846CB1
7475860	24	7475860CD1	78	7475860CB1
7950941	25	7950941CD1	79	7950941CB1
7485334	26	7485334CD1	80	7485334CB1
7220001	27	7220001CD1	81	7220001CB1
5956275	28	5956275CD1	82	5956275CB1
346472	29	346472CD1	83	346472CB1

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
643526	30	643526CD1	84	643526CB1
1483418	31	1483418CD1	85	1483418CB1
2683477	32	2683477CD1	86	2683477CB1
5580991	33	5580991CD1	87	5580991CB1
5605931	34	5605931CD1	88	5605931CB1
6975241	35	6975241CD1	89	6975241CB1
6988529	36	6988529CD1	90	6988529CB1
6996808	37	6996808CD1	91	6996808CB1
7472689	38	7472689CD1	92	7472689CB1
876751	39	876751CD1	93	876751CB1
2512510	40	2512510CD1	94	2512510CB1
7486326	41	7486326CD1	95	7486326CB1
1221545	42	1221545CD1	96	1221545CB1
124737	43	124737CD1	97	124737CB1
1510784	44	1510784CD1	98	1510784CB1
1901257	45	1901257CD1	99	1901257CB1
2044370	46	2044370CD1	100	2044370CB1
2820933	47	2820933CD1	101	2820933CB1
2902793	48	2902793CD1	102	2902793CB1
7486536	49	7486536CD1	103	7486536CB1
8137305	50	8137305CD1	104	8137305CB1
3793128	51	3793128CD1	105	3793128CB1
4001243	52	4001243CD1	106	4001243CB1
6986717	53	6986717CD1	107	6986717CB1
7503512	54	7503512CD1	108	7503512CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
1	95765CD1	g190183	2.6E-76	[Homo sapiens] opiomelanocortin
2	6399886CD1	g6996429	5.8E-262	Krude, H. et al. (1998) Nature Genet. 19:155-157 dJ568C11.3 (novel AMP-binding enzyme similar to acetyl-coenzyme A synthetase (acetate-coA ligase)) [Homo sapiens]
3	6024420CD1	g163497	1.1E-70	Luong, A. et al., (2000) J. Biol. Chem. 275:26458-26466 [Bos taurus] PDI (E.C.5.3.4.1) (protein disulfide isomerase) Yamauchi, K., et al. (1987) Biochem. Biophys. Res. Commun. 146, 1485-1492
4	7481067CD1	g180256	0	[Homo sapiens] preceruloplasmin (EC 1.16.3.1) Waggoner, D.J. et al. (1999) Neurobiol. Dis. 9:221-230; Hellman, N.E. et al. (2000) Gut 47:858-860
5	3378720CD1	g9859003	9.1E-65	[Homo sapiens] tumor metastasis-suppressor
16	1758413CD1	g183178	5.6E-106	[Homo sapiens] hGH-V2
17	7011042CD1	g184808	3.0E-11	Cooke, N.E. et al. (1988) J. Biol. Chem. 263:9001-9006 [Homo sapiens] insulin-like growth factor binding protein complex, acid-labile subunit
18	7427362CD1	g14530679	7.0E-70	Leong, S.R. et al. (1992) Mol. Endocrinol. 6:870-876 WNT3A [Homo sapiens]
19	7485304CD1	g2623871	2.1E-115	Saitoh, T., et al (2001) Biochem. Biophys. Res. Commun. 284:1168-1175 [Gallus gallus] Wnt-14 protein
20	1422394CD1	g15430244	0	Bergstein, J. et al. (1997) Genomics 46:450-458 N-acetylglucosamine-6-sulfatase [Coturnix coturnix]
22	7473674CD1	g13620917	9.0E-46	Dhoot, G.K., et al (2001) Science 293:1663-6. mitochondrial ribosomal protein bMRP63 [Mus musculus]
26	7485334CD1	g10566471	1.2E-73	Suzuki, T., et al (2001) J. Biol. Chem. 276:33181-33195 [Mus musculus] Gliacolin
				Koide, T. et al. (2000) J. Biol. Chem. 275:27957-27963

Table 2

Polypeptide SEQ ID NO:	Incye Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
27	7220001CD1	g11071950	2.2E-115	[Mus musculus] (AB048834) Fcα/m receptor Shibuya, A., et al (2000) Nat. Immunol. 1:441-446
28	5956275CD1	g7259265	4.4E-129	[Mus musculus] contains transmembrane (TM) region Inoue, S., et al. (2000) Biochem. Biophys. Res. Commun. 268, 553-561
39	876751CD1	g15430246	0	nephronectin short isoform [Mus musculus] Brandenberger, R., et al (2001) J. Cell Biol. 154:447-458
40	2512510CD1	g14423349	0	membrane glycoprotein LIG-1 [Homo sapiens]
41	7486326CD1	g3822218	0	[Homo sapiens] chordin Pappano, W. N. et al. (1998) Genomics 52:236-239
48	2902793CD1	g15026974	2.0E-36	obscurin [Homo sapiens] Young, P., et al (2001) J. Cell Biol. 154: 123-136
54	7503512CD1	g14423349	0	membrane glycoprotein LIG-1 [Homo sapiens]

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	095765CD1	235	S92 S211 Y189	N91	signal_cleavage: M1-E23	SPSCAN
					Signal Peptide:	HMMER
					M1-E23, M1-G26, M1-V24, M1-C28, M1-E30	
					Corticotropin ACTH domain:	HMMER_PFAM
					S106-F144, P188-T220	
					Transmembrane Domain: C6-W27	TMAP
					N-terminus is non-cytosolic	
					Pro-opiomelanocortin signature	BLIMPS_PRINTS
					PR00383: Y107-P117, V118-E133, D134-E143, A181-W196, W196-M209	
					PRECURSOR CORTICOTROPIN LIPOTROPIN	BLAST_PRODROM
					PRO-OPIMELANOCORTIN POMC HORMONE	
					SIGNAL ENDORPHIN CLEAVAGE ON PAIR	
					PD004218:L12-Q68	
					PRO-OPIMELANOTROPIN POM PRECURSOR	BLAST_PRODROM
					SIGNAL	
					PD116389: R101-N229, C28-W84	
					HORMONE PRECURSOR CORTICOTROPIN	BLAST_PRODROM
					LIPOTROPIN PRO-OPIMELANOCORTIN POMC	
					CLEAVAGE ON PAIR OF BASIC	
					PD003250: S106-F144	
					CORTICOTROPIN LIPOTROPIN PRECURSOR	BLAST_PRODROM
					PRO-OPIMELANOCORTIN POMC ENDORPHIN	
					HORMONE CLEAVAGE ON PAIR OF	
					PD029102: D164-N229	
					CORTICOTROPIN-LIPOTROPIN	BLAST_DOMO
					DM01793[P01190]1-83: M1-W84	

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1 (cont.)					CORTICOTROPIN-LIPOTROPIN DM00964 P01190 I71-264: K145-E235	BLAST_DOMO
					CORTICOTROPIN-LIPOTROPIN DM01793 P19402 I-79: M1-V78	BLAST_DOMO
					CORTICOTROPIN-LIPOTROPIN DM00964 PN0130 I3-90: A167-E235	BLAST_DOMO
2	6399886CD1	689	S16 S409 S415 S486 S601 S640 T89 T97 T142 T150 T156 T318 T438 T546 T663 T664	N109 N492 N554	Signal Peptide: M1-A25	HMMER
					AMP-binding enzyme: T142-V580	HMMER_PFAM
					Transmembrane Domains: P173-A201, E463-P478 N-terminus is cytosolic	TMAP
					Putative AMP-binding domain signature: A272-V322	PROFILESCAN
					Putative AMP-binding domain BL00455:Y293-Q308	BLIMPS_BLOCKS
					AMP-binding signature PR00154:E286-S297, T298-H306	BLIMPS_PRINTS
					LIGASE SYNTHETASE PROTEIN ENZYME BIOSYNTHESIS ANTIBIOTIC PHOSPHOPANTHETHEINE MULTIFUNCTIONAL REPEAT ACYL-COA PD0000070: T142-T438, D482-I581	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
2 (cont.)					SYNTHETASE LIGASE ACETYL COENZYME A ENZYME ACETATE-COA ACETYL-COA PROTEIN ACYL ACTIVATING PD009307: Y58-L146 PUTATIVE AMP-BINDING DOMAIN DM00073 S46276 96-631: L115-R649 P27550 82-615: L115-R649 S52154 15-546: I141-L648 P16928 102-634: L115-R649 Putative AMP-binding domain signature M291-K302	BLAST_PRODOR          MOTIFS
3	6024420CD1	584	S20 S71 S134 S153 S358 S376 S417 S546 S558 T128 T241 T329 T448 T481 T495 T557 Y356	N58 N160 N238 N340 N370 N436 N540	signal_cleavage: M1-S20	SPSCAN
					Signal Peptide: M1-S20; M1-E23, M1-A26 Thioredoxin domain: Q386-D451 Transmembrane Domain: P199-H216 N-terminal is non-cytosolic	HMMER HMMER_PFAM TMAP
					Thioredoxin family active site: L392-I440 Thioredoxin family proteins signature BL00194; F409-K421	PROFILESAN BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
3 (cont.)					ISOMERASE PRECURSOR PROTEIN SIGNAL REDOX ACTIVE CENTER DISULFIDE ENDOPLASMIC RETICULUM REPEAT PD001708: Q55-L367	BLAST_PRODUM
					REDOX ACTIVE CENTER PROTEIN ISOMERASE PRECURSOR THIOREDOXIN SIGNAL DISULFIDE ENDOPLASMIC ELECTRON PD000175: V389-K496	BLAST_PRODUM
					PROTEIN DISULFIDE-ISOMERASE DM00799 P09102 112-347: L148-D385 DM00799 P54399 129-364: E167-D385 DM00799 A54757 127-358: L148-D385	BLAST_DOMO
					THIOREDOXIN FAMILY DM00054 P09102 349-452: V389-L489	BLAST_DOMO
4	7481067CD1	1049	S38 S180 S221 S241 S317 S318 S383 S551 S604 S638 S736 T155 T184 T198 T203 T208 T222 T311 T389 T414 T430 T452 T515 T570 T589 T661 T688 T783 T855 T1007 T1037 Y134 Y483 Y566 Y715	N111 N137 N169 N245 N270 N392 N460 N475 N549 N575 N781 N829 N830 N908	signal_cleavage: M1-A19	SPSCAN
					Signal Peptide: M1-P16, M1-W18, M1-A19	HMMER



Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
4 (cont.)					Transmembrane Domain: I794-I809	TMAP
					Multicopper oxidase: L900-R1041, L567-C705, L218-C352	HMMER_PFAM
					Multicopper oxidases signatures multicopper_oxidase1.prf: D310-V366, D660-E742, D996-L1050	PROFILES CAN
					multicopper_oxidase2.prf: P1001-N1048	
					Multicopper oxidases signature 1: G680-Y700, G1016-Y1036	MOTIFS
					Multicopper oxidases signature BL00079A: G93-N110, R817-T828, L975-F985, D1014-M1032	BLIMPS_BLOCKS
					FERROXIDASE REPEAT DM00956 P00450 90-336: L89-L334, L438-L684, P838-L1019	BLAST_DOMO
					DM00956 P00450 445-697: H437-L684, L89-S318, Y819-W1018,	
					H791-S831 DM00956 P00450 804-1038: H791-Y1021, H437-G680, L89-L334	
					DM00956 P12259 83-308: G90-P249, G439-S577, I794-V992,	
					E252-L334, D595-T675, S538-L552	
					FACTOR PRECURSOR GLYCOPROTEIN PLASMA REPEAT SIGNAL COAGULATION BLOOD VIII CALCIUM PD002090: H368-K559, R22-N205, K707-P893	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
4 (cont.)					ATP/GTP-binding site motif A (P-loop): G335-S342	MOTIFS
5	3378720CD1	383	S23 S90 S119 S247 S334 S340 T4 T19 T73 T78 T81 T109 Y345	N107	signal_cleavage: M1-E57	SPSCAN
					Transmembrane Domains: H38-A61, C136-Y160, Y181-F198, K201-R229, N261-F281, H294-L314 N terminus is non-cytosolic	TMAP
					PROTEIN TRANSMEMBRANE LONGEVITY ASSURANCE FACTOR UOG1 SIMILAR S CEREVISIAE PD006418: S119-L369	BLAST_PRODOM
6	938824CD1	72			signal_cleavage: M1-A63	SPSCAN
					Signal Peptide: M1-G29	HMMER
					Transmembrane Domain: P9-A36	TMAP
					N terminus is non-cytosolic	
					Galactokinase signature galactokinase, prf: S17-A67	PROFILESCAN
7	1683721CD1	91	S41 T34 T53		signal_cleavage: M1-P23	SPSCAN
					Signal Peptide: M1-P23	HMMER
8	1694122CD1	160	T19		signal_cleavage: M1-R62	SPSCAN
9	1970615CD1	95	S55		signal_cleavage: M1-A26	SPSCAN
					Signal Peptide: M1-A26, M1-R27, M1-V28, M1-C30	HMMER
					Transmembrane Domain: Q4-C20 N-terminus is non-cytosolic	TMAP

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
9 (cont.)					Eukaryotic cobalamin-binding proteins signature cobalamin_binding.prf: D25-R72	PROFILES CAN
10	2314152CD1	92	S11 S23 S33 S52 T65 Y43		signal_cleavage: M1-S19	SPSCAN
					Signal Peptide: M1-R25, M1-S24	HMMER
					Endopeptidase Clp active sites clpp_protease_ser.prf: L32-G80	PROFILES CAN
11	2886225CD1	71	S41 T37 Y67		signal_cleavage: M1-L29	SPSCAN
					Signal Peptide: M1-L29	HMMER
					Transmembrane Domain: H15-C33	TMAP
					N-terminus is non-cytosolic	
					Vitamin K-dependent carboxylation domain glu_carboxylation.prf: M1-G70	PROFILES CAN
12	6144418CD1	100	S46 S69 T50 T60 T84	N67 N85	signal_cleavage: M1-G36	SPSCAN
					Signal Peptide: M1-A29, M1-V31	HMMER
					Transmembrane Domain: L8-G36	TMAP
					N-terminus is non-cytosolic	
13	6834184CD1	122		N108	signal_cleavage: M1-E38	SPSCAN
					Signal Peptide: M1-G25	HMMER
					Transmembrane Domain: G8-T36, C81-I105	TMAP
					N-terminus is non-cytosolic	
14	6951005CD1	113	S25 T35 Y81		signal_cleavage: M1-D28	SPSCAN
					Signal Peptide: M1-Q29, M1-D28	HMMER
					Transmembrane Domain: S50-S72	TMAP
					N-terminus is non-cytosolic	
15	7250331CD1	85			signal_cleavage: M1-L27	SPSCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15 (cont.)					Signal Peptide: M1-P23, M1-V25, M1-R26, M1-L27, M1-C35 Immunoglobulins and major histocompatibility complex proteins signature ig_mhc.prf: L14-L64	HMMER PROFILESAN
16	1758413CD1	256	S88 S97 S111 S132 S181 S194 T53		signal_cleavage: M1-A26	SPSCAN
					Signal Peptide: M1-A26	HMMER
					Transmembrane Domain: T7-L35	TMAP
					N-terminus is non-cytosolic	
					Somatotropin hormone family: L9-M151	HMMER_PFAM
					Somatotropin, prolactin and related hormones	PROFILESAN
					somatotropin_1.prf: S88-H138	
					Somatotropin, prolactin and related hormones	MOTIFS
					signature 1: C79-W112	
					Somatotropin, prolactin and related hormones	BLIMPS_BLOCKS
					BL00266: L35-Y61, C79-V116, D135-M151, P201-P224	
					Somatotropin hormone family signature PR00836: C79-E92, L101-L119	BLIMPS_PRINTS
					GROWTH HORMONE VARIANT II PRECURSOR	BLAST_PRODROM
					GHV2 PLACENTA SIGNAL ALTERNATIVE	
					SPLICING	
					PD084405: P165-V256	
					HORMONE PRECURSOR SIGNAL PITUITARY	BLAST_PRODROM
					GROWTH SOMATOTROPIN PROLACTIN	
					GLYCOPROTEIN PRL	
					PD000259: T7-M151	

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
16 (cont.)					SOMATOTROPIN, PROLACTIN AND RELATED HORMONES DM00125 P09587 17-227: C17-P228 DM00125 P01243 17-212: C17-G163 DM00125 I67409 17-212: C17-G163 DM00125 P01242 17-212: C17-G163	BLAST_DOMO
17	7011042CD1	287	S16 S46 S91 S147 T76 T105 T214 T242	N42 N176	signal_cleavage: M1-A18	SPSCAN
					Signal Peptide: M1-A18, M1-W17	HMMER
					Leucine Rich Repeat: N254-P278, Q100-A122, A171-D199, T205-P228, G123-S146, T76-E99, S147-P170, K229-P253	HMMER_PFAM
					Leucine Rich Repeat PR00019A: L124-L137	BLIMPS_PRINTS
					Leucine zipper pattern: L59-L80, L66-L87	MOTIFS
18	7427362CD1	366	S89 T34 T81 T150 T215 T226 T257 T356 T363	N103	signal_cleavage: M1-A29	SPSCAN
					Signal Peptide: M1-A29	HMMER
					Transmembrane Domain: P6-L33	TMAP
					N-terminus is non-cytosolic	
					wnt family of developmental signaling proteins: A58-K365	HMMER_PFAM
					Wnt-1 family signature wnt1.prf: M196-K245	PROFILES SCAN
					Wnt-1 family signature: C216-C225	MOTIFS
					Wnt-1 family proteins BL00246: M196-Y248, N319-C364, A85-C104, G118-D152, W163-E187	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
18 (cont.)					DEVELOPMENTAL GLYCOPROTEIN PRECURSOR SIGNAL WNT1 WNT5A WNT2 EXTRACELLULAR MATRIX PD000810: C59-A195, E153-K365	BLAST_PRODOM
					WNT-1 FAMILY DM00403 P27467 22-351: K67-K365 DM00403 P21551 32-368: L64-C364 DM00403 P28465 13-351: C59-K365 DM00403 P22727 24-363: A154-C364	BLAST_DOMO
19	7485304CD1	416	S72 S120 S228 S238 T87 T136 T203 T268 T413 Y342	N158	wnt family of developmental signaling proteins: Q113-K415	HMMER_PFAM
					Wnt-1 family signature wnt1.prf: D250-K298	PROFILES CAN
					Wnt-1 family signature: C269-C278	MOTIFS
					Wnt-1 family proteins BL00246: A140-C159, G171-D205, W216-R240, A249-Y301, S369-C414	BLIMPS_BLOCKS
					PROTEIN DEVELOPMENTAL GLYCOPROTEIN PRECURSOR SIGNAL WNT1 WNT5A WNT2 EXTRACELLULAR MATRIX PD000810: L119-A251, D206-K415	BLAST_PRODOM
					WNT-1 FAMILY DM00403 P49340 30-391: L119-S365, A334-C414, S69-Q106 DM00403 P21551 32-368: L119-C414 DM00403 P47793 24-351: C114-K415 DM00403 P22727 24-363: E211-L410, K118-G209	BLAST_DOMO

Table 3

SEQ ID NO:	Incye Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
20	1422394CD1	871	S27 S206 S288 S425 S468 S488 S505 S516 S520 S635 T24 T66 T96 T107 T367 T376 T400 T452 T484 T530 T611 T615 T781 Y645	N64 N111 N131 N148 N170 N197 N240 N623 N773 N783	signal_cleavage: M1-C22	SPSCAN
					Signal Peptide: M1-C22	HMIMER
					Transmembrane Domain: C5-L21, S353-D370	TMAP
					N-terminus is cytosolic	
					Sulfatase: P43-T467	HMIMER_PFAM
					Sulfatases signature 1: P85-G97	MOTIFS
					Sulfatases proteins	BLIMPS_BLOCKS
					BL00523: P43-S59, C87-K98, G134-L144, P214-H225, V289-G318, D363-G373, Y800-Q809	
					ARYLSULFATASE SIGNAL GLYCOPROTEIN LYSOSOME SULPHOHYDROLASE MUCOPOLYSACCHARIDOSIS PD001700: P43-I200, Q279-P378, K330-E392	BLAST_PRODROM
					N-ACETYLGLUCOSAMINE-6-SULFATASE PRECURSOR GLUCOSAMINE-6-SULFATASE HYDROLASE LYSOSOME SIGNAL GLYCOPROTEIN MUCOPOLYSACCHARIDOSIS PD0022780: C766-G837	BLAST_PRODROM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
20					SIMILAR TO SULFATASE	BLAST_PRODOM
(cont.)					PD122645: V385-Q487, H700-F765	
					ARYLSULFATASE; SULFATASE; PLANT;	BLAST_DOMO
					DM08669 Q10723 23-520: R280-R490, R42-Q279, N773-P843	
					DM08669 P14217 24-519: L278-G373, R42-P269, S770-P843, E294-Y310	
					DM01026 P08842 24-548: V289-D390, R42-Y143	
					DM01026 P50473 63-522: E273-L389, R42-P153	
21	1336022CD1	100	S3 S72 S93	N5	signal_cleavage: M1-V24	SPSCAN
					Signal Peptide: M1-V24	HMMER
					Transmembrane Domain: V4-P23	TMAP
					N-terminus is non-cytosolic	
22	7473674CD1	102	S28 S102 T79 T98 T138 Y143	N96	signal_cleavage: M1-G19	SPSCAN
23	7475846CD1	117			signal_cleavage: M1-C32	SPSCAN
24	7475860CD1	150	S65 T58 T99 T113 T138 Y143	N46 N66	signal_cleavage: M1-A34	SPSCAN
25	7950941CD1	89	S10 S34 S39 S49		Signal Peptide: M1-G24	HMMER
26	7485334CD1	287	S111 S167 S255 T28 T142 T192		signal_cleavage: M1-A15	SPSCAN
					Signal Peptide: M1-A20, M1-A21	HMMER
					Clq domain: A160-L284	HMMER_PFAM
					Clq domain signature: F177-Y207	MOTIFS



Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
26 (cont.)					C1q domain proteins BL01113: V174-I209, D243-K262, S277-P286, G85-S111	BLIMPS_BLOCKS
					Complement C1Q domain signature PR00007: P168-K194, F195-G214, D243-D264, K275-Y285	BLIMPS_PRINTS
					Collagen triple helix repeat (20 copies): P71-V129	HMMER_PFAM
					Lysosome-associated membrane glycoproteins lamp_2.prf: E145-L175	PROFILES SCAN
					PRECURSOR SIGNAL COLLAGEN REPEAT HYDROXYLATION GLYCOPROTEIN CHAIN PLASMA EXTRACELLULAR MATRIX PD002992: P168-L284	BLAST_PRODROM
					COLLAGEN ALPHA PRECURSOR REPEAT SIGNAL CONNECTIVE TISSUE EXTRACELLULAR MATRIX PD000007: G43-G118	BLAST_PRODROM
					PROCOLLAGEN TYPE XVII ALPHA 1 BULLOUS PEMPHIGOID AUTO-ANTIGEN CELL ADHESION PD071338: G43-G112	BLAST_PRODROM
					PRECOLLAGEN P PRECURSOR SIGNAL PD072959: G43-G128	BLAST_PRODROM
					C1Q DOMAIN DM00777 S23297 465-674: P84-L283 DM00777 P23206 477-673: G88-P286 DM00777 P98085 222-418: G85-D287 DM00777 P27658 551-743: G88-P286	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
27	7220001CD1	578	S39 S108 S189 S252 S297 S302 S406 S483 S494 S526 T6 T38 T88 T234 T272 T336 T350 T351 T438 T487 T525 T570 Y24	N212 N322	signal_cleavage: M1-A61	SPSCAN
					Signal Peptide: M46-A61, M46-P63, M46-Q64	HMMER
					Immunoglobulin domain: G120-I200	HMMER_PFAM
					Transmembrane Domain: S39-P67 R496-R518	TMAP
					N-terminus is cytosolic	
					POLYMERIC IMMUNOGLOBULIN RECEPTOR PRECURSOR PLGR CONTAINS: SECRETORY COMPONENT IMMUNOGLOBULIN FOLD REPEAT PD003917: G120-E203 (P-value = 5.4e-09)	BLAST_PRODOM
					IMMUNOGLOBULIN DM00001 P01833 41-120: H128-G201 DM00001 P15083 41-120: H128-F208 DM00001 P01832 28-125: G120-G201 DM00001 S4884 41-120: H128-G201	BLAST_DOMO
28	5956275CD1	285	S109 S133 S256 T38 T91 T100 T191 Y125		signal_cleavage: M1-A28	SPSCAN
					Signal Peptide: M1-A28	HMMER
					Immunoglobulin domain: G42-V129	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
28 (cont.)					Transmembrane Domain: Q3-T31 N-terminus is non-cytosolic	TMAP
29	346472CD1	72	S25		Cell attachment sequence: R197-D199 signal_cleavage: M1-C14 Signal Peptide: M1-S16, M1-S20, M1-N21, M1-F22, M1-H24, M1-S25, M1-K27 Aminotransferases class-V pyridoxal-phosphate attachment site: S25-S72	MOTIFS SPSCAN HMMER PROFILES SCAN
30	643526CD1	72			signal_cleavage: M1-C18 Signal Peptide: M1-C18, M1-S19, M1-E20, M1-S21, M1-G23, M1-S24, M1-P26 Transmembrane Domain: S24-V41	SPSCAN HMMER TMAP
31	1483418CD1	149	S65 S70		signal_cleavage: M1-H29 Signal Peptide: M1-G32, M1-V34, M1-S35 Transmembrane Domain: G32-P53 N-terminus is non-cytosolic	SPSCAN HMMER TMAP
32	2683477CD1	100	S71 T54 T81	N69	Signal Peptide: M7-L35 Transmembrane Domain: S13-R38 N-terminus is non-cytosolic	HMMER TMAP
33	5580991CD1	78	S50		signal_cleavage: M1-C24 Signal Peptide: M4-S22, M4-C24, M4-P25, M4-S26, M4-S29 Transmembrane Domain: F13-F41 N-terminus is non-cytosolic	SPSCAN HMMER TMAP
34	5605931CD1	75	S65		CBF/NF- $\gamma$ subunits signatures: M4-S73 Signal Peptide: M29-A43 Signal Peptide: M1-Q28	PROFILES SCAN HMMER HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
34					Transmembrane Domain: A5-F25 N-terminus is non-cytosolic	TMAP
(cont.)						
35	6975241CD1	111	S6 S51 S75		signal_cleavage: M1-E22 Signal Peptide: M1-E22, M1-K24, M1-K24, M1-G28, M1-I30	SPSCAN HMMER
36	6988529CD1	72	S30 T35		signal_cleavage: M1-A15 Signal Peptide: M1-A15, M1-T17, M1-A20	SPSCAN HMMER
37	6996808CD1	90	S34 S42 S68		signal_cleavage: M1-C45 Signal Peptide: M26-C45, M17-L38, M17-S41 Transmembrane Domains: Q12-F40, R44-I72 N-terminus is non-cytosolic	SPSCAN HMMER TMAP
38	7472689CD1	283	S66 S140 S197 T259	N131	Signal Peptide: M21-G38	HMMER
39	876751CD1	566	S111 S238 S407 S485 S556 T167 T176 T308 T312 T316 T320 T346 T527 Y206	N274	Signal Peptide: M1-A19	SPSCAN
					Signal Peptide: M1-A18, M1-E20 EGF-like domain: C94-C128, C174-C213, C219-C254, C134-C168, C61-C87 MAM domain: C423-E564 Anaphylatoxin domain proteins BL01177: R99-L117, P163-S180 Calcium-binding EGF-like domain BL01187: C105-Y120, C168-A179	HMMER HMMER_PFAM HMMER_PFAM BLIMPS_BLOCKS BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
39					PRECURSOR GLYCOPROTEIN SIGNAL TRANSMEMBRANE HYDROLASE PROTEIN REPEAT RECEPTOR PHOSPHATASE NEUROPILIN PD001482: W432-E564	BLAST_PRODOM
(cont.)					PROLINE-RICH PROTEIN 3 DM00215[P41479]30-111: P297-I369	BLAST_DOMO
					Cell attachment sequence: R383-D385	MOTIFS
					Aspartic acid and asparagine hydroxylation site: C105-C116 C187-C198 C232-C243	MOTIFS
					EGF-like domain signature 1: C76-C87	MOTIFS
					EGF-like domain signature 2: C76-C87 C114-C128 C241-C254	MOTIFS
					Calcium-binding EGF-like domain pattern signature: D90-C114 D170-C196 D215-C241	MOTIFS
40	2512510CD1	1093	S81 S270 S326 S356 S379 S403 S473 S588 S705 S734 S822 S850 S970 S987 S998 T192 T214 T320 T370 T499 T538 T604 T609 T688 T736 T771 T818 T828 T848 T936 Y670	N74 N150 N246 N292 N318 N684	Signal_Peptide: M1-A33	SPSCAN
					Signal Peptide: M1-A33, M1-A34	HMMER

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
40 (cont.)					Leucine Rich Repeat: S332-R355, S236-S259, K308-S331, R212-N235, G407-K430, S356-D382, N93-S114, S189-P211, K260-T283, S383-E406, P164-S187, S140-P163, W69-P92, A284-Q307, H116-K136	HMMER_PFAM
					Leucine rich repeat C-terminal domain: D440-D490	HMMER_PFAM
					Immunoglobulin domain: G707-M765, T613-A674, G509-I579	HMMER_PFAM
					Transmembrane Domain: R13-A33	TMAP
					N terminus non-cytosolic	BLIMPS_PRINTS
					Bee Venom Hyaluronidase PR00847F: R738-V754	BLIMPS_PRINTS
					Leucine-rich repeat signature PR00019 L91-L104, L141-V154	BLAST_PRODROM
					MEMBRANE GLYCOPROTEIN PD129774: M765-S1093 PD172109: D491-F583 PD165826: E29-T70	BLAST_PRODROM
					SIMILARITY MULTIPLE LEUCINE RICH PD037237: L432-I610	BLAST_PRODROM
					IMMUNOGLOBULIN DM000001 P35918 668-745: L699-A774 DM000001 A46065 668-745: L699-A774	BLAST_DOMO
					Leucine zipper pattern: L52-L73 L59-L80	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
41	7486326CD1	915	S38 S98 S130 S132 S151 S153 S195 S201 S245 S306 S353 S488 S596 S751 S882 T304 T367 T592 T775 T859 T902 T904 Y152 Y745	N217 N351 N365 N434	Signal_Peptide: M1-P23	SPSCAN
					Signal Peptide: M1-P23, M1-G26	HMIMER
					von Willebrand factor (growth regulator) type C domain: C51-C125, C832-C892, C744-C810, C665-C722	HMIMER_PFAM
					Transmembrane Domain: P5-R25	TMAP
					N terminus non-cytosolic	
					CHORDIN	BLAST_PRODOR
					PD018424: Q440-L618, G232-Y439, F141-P292, A413-V605	
					PD069130: P811-C892PD015143: P662-D727	
					VON WILLEBRAND FACTOR TYPE C REPEAT	BLAST_DOMO
					DM00551 A55195 22-117: P31-P126	
					DM00551 A55195 752-835: E724-K808, P656-C719	
					DM00551 A55195 637-751: R660-C722, R738-C810, T592-L618	
					DM00551 A55195  836-920: Q809-C892	
					Cell attachment sequence R165-D167	MOTIFS
					VWFC domain signature:	MOTIFS
					C686-C722 C771-C810	

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
41					Leucine zipper pattern: L315-L336	MOTIFS
(cont.)						
42	1221545CD1	113	S66 S71 T44 T100	N88	Signal_Peptide: M1-T45	SPSCAN
					Signal Peptide: M1-D18, M1-G20	HMMER
					Transmembrane Domain: T4-R24	TMAP
					N terminus cytosolic	
43	124737CD1	91	S30 S43		Signal_Peptide: M1-A26	SPSCAN
					Signal Peptide: M1-A26, M1-L31	HMMER
44	1510784CD1	83	S33 S61 S71		Signal_Peptide: M1-S61	SPSCAN
					Signal Peptide: M1-H17, M3-P23	HMMER
					Defensin signature: PR00217A L58-C67	BLIMPS_PRINTS
45	1901257CD1	128	S28 T31 T38		Signal_Peptide: M1-S27	SPSCAN
					Signal Peptide: M7-S34, M46-S74	HMMER
					Transmembrane Domain: A4-A25	TMAP
					N terminus non-cytosolic	
46	2044370CD1	84	S31 S49 T22	N38	Signal_Peptide: M1-G21	SPSCAN
					Signal Peptide: M1-G21	HMMER
					Transmembrane Domain: L3-Y23	TMAP
					Sushi (SCR complement) domain: PF00084A:C69-P73	BLIMPS_PFAM
47	2820933CD1	109	S33 S52 S97		Signal_Peptide: M1-S21	SPSCAN
					Signal Peptide: M1-S21, M1-E31	HMMER
48	2902793CD1	159	S85 S154 T122		Signal_Peptide: M1-A23	SPSCAN
					Signal Peptide: M1-C25	HMMER
					Immunoglobulin domain: G54-C112	HMMER_PFAM
					Transmembrane Domain: V4-S21	TMAP
					N terminus non-cytosolic	



Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
49	7486536CD1	242	S86 S100 S104 S115 S139 S229 T30 T134 Y176	N76	Signal Peptide: M1-L19, M1-N21	HM/MER
					Transmembrane Domain: T6-L29 N terminus non-cytosolic	TMAP
50	8137305CD1	542	S74, S90, S95, S111, S126, T225, Y245, T231, S320, S338, S379, S424, T536	N41, N333	Signal cleavage: M1-F28	SPSCAN
					Signal Peptide: M1-C34	HM/MER
					Transmembrane Domain: I7-L32; N-terminus is non-cytosolic	TMAP
51	3793128CD1	105	S30, T37	N55	Signal cleavage: M1-R22	SPSCAN
					Signal Peptide: M1-S19	HM/MER
					Transmembrane Domain: P7-Y34; N-terminus is non-cytosolic	TMAP
52	4001243CD1	102			Signal Peptide: M76-A95, M1-A27	HM/MER
53	6986717CD1	129	S31, S77, T125	N83	Signal Peptide: M25-R48, M25-A53, M25-Q51	HM/MER
					Transmembrane Domain: P4-R19; N-terminus is non-cytosolic	TMAP

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
54	7503512CD1	1070	S81 S247 S303 S333 S356 S380 S450 S565 S682 S711 S799 S827 S947 S964 S975 T192 T297 T347 T476 T515 T581 T586 T665 T713 T748 T795 T805 T825 T913 Y647	N74 N150 N223 N269 N295 N661	signal_cleavage: M1-A33	SPSCAN
					Signal Peptide: M1-A33, M1-A34	HMIMER
					Leucine Rich Repeat: S309-R332, S213-S236, K285-S308, S333-S356, G384-K407, N93-S114, W69-P92, K237-T260, S360-E383, P164-S187, S140-P163, S189-N212, A261-Q284, H116-K136	HMIMER_PFAM
					Leucine rich repeat C-terminal domain: D417-D467	HMIMER_PFAM
					Leucine rich repeat N-terminal domain: P40-P67	HMIMER_PFAM
					Immunoglobulin domain: G684-M742, T590-A651, G486-I556	HMIMER_PFAM
					Cytosolic domain: Y793-S1070	TMHMMER
					Transmembrane domain: V770-I792	
					Non-cytosolic domain: M1-T769	
					RECEPTOR INTERLEUKIN-1 P	BLIMPS_PRODUM
					PD02870: F634-V666, L725-P759	

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
54 (cont.)					MEMBRANE GLYCOPROTEIN PD129774: M742-S1070 PD172109: D468-F560 PD165826: E29-T70	BLAST_PRODOM
					KEK1 PRECURSOR T2ID12.9 SPLICING SIGNAL ALTERNATIVE KEK2 PD037237: L409-I587	BLAST_PRODOM
					IMMUNOGLOBULIN DM00001 P35918 668-745: L676-A751 DM00001 A46065 668-745: L676-A751	BLAST_DOMO
					Leucine zipper pattern: L52-L73, L59-L80	MOTIFS

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
55/095765CBI/1315	1-68, 1-341, 1-346, 1-359, 1-365, 1-367, 1-435, 1-566, 1-589, 1-625, 1-647, 2-367, 3-365, 3-367, 4-367, 5-367, 6-367, 7-68, 7-367, 9-367, 11-367, 67-367, 154-367, 244-367, 290-358, 361-547, 361-550, 362-957, 363-384, 363-387, 363-394, 363-399, 363-400, 363-403, 363-411, 363-417, 363-418, 363-424, 363-434, 363-436, 363-443, 363-444, 363-446, 363-450, 363-454, 363-459, 363-460, 363-461, 363-464, 363-467, 363-469, 363-473, 363-476, 363-477, 363-480, 363-484, 363-487, 363-489, 363-490, 363-491, 363-492, 363-496, 363-497, 363-498, 363-500, 363-501, 363-504, 363-509, 363-510, 363-511, 363-512, 363-515, 363-517, 363-518, 363-521, 363-526, 363-527, 363-528, 363-529, 363-530, 363-532, 363-533, 363-534, 363-536, 363-538, 363-539, 363-540, 363-541, 363-544, 363-545, 363-547, 363-549, 363-550, 363-551, 363-553, 363-554, 363-555, 363-556, 363-559, 363-560, 363-562, 363-563, 363-564, 363-565, 363-566, 363-567, 363-569, 363-571, 363-572, 363-574, 363-580, 363-582, 363-584, 363-586, 363-590, 363-592, 363-594, 363-595, 363-600, 363-604, 363-608, 363-611, 363-630, 363-925, 363-950, 363-956, 363-957, 365-447, 365-502, 365-510, 365-544, 365-550, 365-554, 365-560, 365-948, 367-443, 367-955, 379-956, 381-956, 388-957, 390-936, 394-956, 401-957, 414-955, 416-950, 421-956, 425-956, 433-554, 442-957, 444-957, 448-957, 451-956, 451-957, 452-955, 456-956, 458-953, 458-955, 461-956, 472-952, 473-957, 477-956, 479-956, 479-957, 484-956, 485-615, 487-956, 491-936, 493-957, 494-956, 495-956, 496-957, 499-954, 499-956, 499-957, 505-956, 505-957, 509-957, 512-949, 516-955, 529-956, 531-957, 537-956, 539-956, 539-957, 541-957, 545-955, 548-947, 548-956, 553-956, 555-953, 563-778, 563-912, 563-957, 564-955, 568-957, 569-965, 572-965, 580-957, 584-1206, 584-1315, 585-957, 586-949, 590-956, 593-956, 593-957, 594-965, 595-734, 595-957, 595-965, 597-743, 601-957, 605-794, 606-710, 607-882, 610-956, 624-954, 629-955, 630-947, 636-957, 641-957, 644-954, 651-921, 655-810, 657-929, 659-777, 659-798, 659-799, 659-940, 664-904, 671-947, 671-957, 678-897, 678-919, 690-835, 705-957, 723-956, 723-957, 729-956, 731-956, 731-957, 775-955, 778-957, 784-957, 805-957, 813-957, 823-957, 836-957, 850-957, 854-957, 881-954, 911-956
56/6399886CBI/3796	1-268, 1-287, 184-830, 186-653, 280-606, 280-753, 280-767, 280-836, 280-852, 280-924, 280-936, 365-765, 365-768, 365-794, 439-1252, 497-916, 571-1252, 621-1286, 645-836, 709-1568, 732-1214, 743-991, 773-1403, 810-1354, 810-1367, 810-1440, 843-1013, 847-1135, 856-1475, 858-1504, 867-1103, 928-1240, 942-1561, 961-1568, 985-1521, 1010-1130, 1010-1451, 1155-1438, 1158-1413, 1207-1495, 1207-1737, 1272-1533, 1272-1729, 1451-1529, 1506-1963, 1531-1871, 1534-1829, 1540-1786, 1543-1737, 1543-1814, 1597-1835, 1649-2212, 1669-2235, 1703-1891, 1711-2149, 1757-2033, 1802-2377, 1883-2433, 1914-2400, 1970-2635, 2001-2230, 2001-2677, 2014-2440, 2036-2249, 2036-2521, 2081-2503, 2086-2350, 2124-2563, 2151-2480, 2177-2453, 2184-2418, 2251-2875, 2257-2524, 2305-2534, 2313-2532, 2313-2938, 2314-3011, 2320-2540, 2331-2799, 2332-2647, 2390-2549, 2415-2612, 2459-2744, 2592-2715, 2602-2830, 2611-2886, 2681-2880, 2681-2928, 2684-2975, 2700-2955, 2710-2998, 2716-3047, 2720-3177, 2725-2947, 2745-2901, 2778-3062, 2792-3019, 2792-3188, 2794-2993, 2808-3112, 2811-3065, 2825-3106, 2826-3084, 2826-3104, 2831-3039, 2868-3094, 2868-3100, 2876-3144, 2883-3112, 2883-3141, 2883-3310, 2904-3133, 2904-3460, 2914-3165, 2951-3227, 2959-3180,

Table 4

Polynucleotide SEQ ID NO:/ Incye ID/ Sequence Length	Sequence Fragments
56 (cont.)	2977-3227, 2991-3245, 3026-3426, 3027-3314, 3115-3376, 3124-3390, 3129-3730, 3181-3406, 3199-3444, 3204-3445, 3217-3786, 3234-3494, 3240-3488, 3281-3492, 3315-3796, 3327-3775, 3341-3580, 3437-3681, 3483-3711, 3484-3738, 3572-3781, 3629-3796
57/6024420CB1/2983	1-186, 1-286, 1-506, 1-517, 1-631, 1-637, 1-711, 1-787, 2-591, 2-638, 2-667, 3-660, 4-689, 14-650, 70-526, 110-705, 202-837, 272-862, 281-825, 284-883, 298-1007, 319-948, 332-906, 334-884, 346-1151, 403-540, 408-1110, 448-1180, 462-729, 529-1218, 569-1233, 578-1254, 607-852, 608-1341, 612-1232, 664-1503, 668-1227, 714-1339, 740-1382, 781-1404, 811-1688, 882-1650, 914-1605, 916-1611, 931-1462, 947-1202, 955-1479, 998-1585, 999-1798, 1019-1566, 1077-1903, 1078-1637, 1081-1664, 1093-1614, 1127-1943, 1130-1810, 1141-1783, 1147-1873, 1179-1881, 1217-1782, 1264-1829, 1267-1889, 1271-1878, 1294-1918, 1317-1930, 1415-1878, 1709-2983, 1740-2008
58/7481067CB1/3840	1-222, 160-947, 250-418, 514-606, 775-1155, 781-861, 1027-1420, 1086-1379, 1321-1662, 1463-2176, 1463-2964, 1664-1856, 2038-2353, 2248-2294, 2250-2759, 2264-2506, 2423-2507, 2458-2696, 2550-2824, 2571-2825, 2584-3076, 2708-2984, 2713-2753, 2824-3128, 2937-3384, 2970-3384, 3059-3570, 3221-3840
59/3378720CB1/1570	1-293, 1-363, 1-404, 1-437, 1-452, 1-511, 1-521, 1-544, 1-553, 1-557, 1-565, 1-569, 6-276, 41-656, 49-720, 68-639, 89-709, 121-747, 215-841, 219-429, 225-755, 249-864, 296-906, 339-955, 380-934, 423-599, 450-750, 461-739, 634-1045, 647-791, 771-1317, 825-995, 841-1425, 875-1481, 911-1528, 911-1536, 978-1535, 984-1515, 994-1402, 1070-1526, 1093-1316, 1130-1570, 1344-1570
60/938824CB1/409	1-245, 1-294, 1-338, 207-409, 333-393
61/1683721CB1/953	1-189, 1-209, 1-439, 1-470, 1-581, 10-695, 26-658, 209-685, 251-573, 258-536, 269-489, 298-734, 321-839, 324-540, 360-521, 386-862, 388-742, 522-948, 556-953, 646-953
62/1694122CB1/890	1-286, 2-25, 2-479, 25-44, 25-311, 25-314, 25-331, 25-337, 25-338, 25-416, 27-331, 29-110, 51-145, 78-554, 94-529, 96-554, 120-546, 160-439, 160-653, 160-705, 161-435, 164-424, 201-860, 301-554, 301-857, 326-628, 335-621, 336-437, 336-554, 362-487, 362-584, 362-862, 362-881, 362-890, 424-554, 428-554, 445-554, 448-554, 452-693, 452-886, 467-554, 482-725, 516-554, 553-774, 553-807, 553-822, 554-829, 555-869
63/1970615CB1/1960	1-421, 84-333, 117-709, 144-700, 178-360, 178-374, 282-525, 285-709, 286-550, 286-852, 311-555, 423-708, 423-709, 583-736, 598-880, 598-881, 598-1013, 598-1015, 598-1074, 598-1094, 598-1120, 634-1049, 653-1211, 670-921, 678-929, 709-968, 719-1226, 728-1117, 731-978, 740-948, 740-952, 740-1006, 742-965, 742-1294, 747-1016, 762-1213, 850-1163, 869-1053, 922-1173, 922-1186, 983-1255, 1038-1282, 1049-1308, 1049-1327, 1050-1546, 1058-1342, 1060-1546, 1095-1337, 1117-1546, 1117-1603, 1182-1395, 1198-1891, 1202-1447, 1202-1749, 1219-1470, 1226-1466, 1287-1516, 1338-1625, 1351-1590, 1355-1652, 1369-1960, 1447-1640, 1447-1887, 1447-1901, 1472-1774, 1474-1733

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
64/2314152CB1/832	1-519, 11-466, 13-433, 29-187, 29-476, 29-486, 29-588, 29-628, 32-408, 34-490, 35-239, 35-348, 35-415, 36-655, 36-662, 49-548, 57-457, 58-643, 61-648, 66-760, 77-603, 91-552, 91-598, 94-644, 172-415, 219-832, 412-686, 501-757, 509-832, 525-774
65/2886225CB1/546	1-454, 273-533, 275-546
66/6144418CB1/890	1-123, 1-596, 42-619, 42-630, 60-709, 130-660, 188-658, 188-732, 222-737, 222-778, 222-890, 255-732, 258-686
67/6834184CB1/807	1-554, 1-687, 1-807
68/6951005CB1/677	1-564, 1-643, 4-677
69/7250331CB1/617	1-287, 3-397, 241-532, 269-598, 289-600, 307-497, 344-498, 366-597, 370-617, 441-586, 441-615
70/1758413CB1/795	1-456, 292-768, 292-795, 417-695, 573-707
71/7011042CB1/1677	1-895, 241-1100, 241-1660, 251-753, 896-1043, 920-1043, 1259-1677
72/7427362CB1/1402	1-1268, 268-527, 270-1268, 787-926, 875-1402
73/7485304CB1/1251	1-1251, 73-171, 255-1081
74/1422394CB1/4961	1-493, 34-650, 38-279, 38-311, 39-650, 48-597, 51-590, 304-484, 304-753, 309-899, 421-957, 488-710, 559-633, 582-971, 632-777, 717-1013, 738-1442, 748-1151, 748-1383, 756-951, 958-1606, 1126-1486, 1126-1815, 1273-1817, 1277-1828, 1291-1949, 1322-1922, 1336-1840, 1336-1859, 1336-1908, 1438-1714, 1442-2028, 1509-1757, 1509-2045, 1549-2251, 1754-2257, 1760-2378, 1905-2232, 1905-2438, 1905-2485, 1905-2491, 1905-2571, 1905-2626, 1908-2337, 1925-2480, 1967-2630, 1969-2201, 2010-2503, 2012-2506, 2034-2531, 2037-2626, 2044-2617, 2067-2282, 2070-2350, 2108-2715, 2117-2622, 2132-2324, 2132-2636, 2157-2816, 2183-2689, 2209-2626, 2244-2715, 2278-2790, 2301-2758, 2363-2644, 2363-2647, 2475-2751, 2479-2794, 2479-2884, 2487-2836, 2497-2791, 2498-2827, 2507-2858, 2680-2901, 2680-3121, 2702-3304, 2723-3251, 2774-3063, 2825-3066, 2830-3456, 2837-3405, 2844-3134, 2859-3128, 2861-3141, 2863-3149, 2871-3479, 3077-3336, 3077-3588, 3141-3409, 3141-3601, 3152-3414, 3182-3457, 3260-3554, 3277-3672, 3299-3596, 3299-3602, 3329-3556, 3329-3566, 3349-3612, 3377-3678, 3390-3613, 3408-3634, 3420-3691, 3506-3778, 3519-3673, 3595-4170, 3602-3798, 3625-3922, 3690-3884, 3697-3928, 3716-3979, 3716-4001, 3862-4072, 3865-4142, 3890-4136, 3967-4157, 4002-4269, 4029-4210, 4068-4311, 4070-4381, 4072-4316, 4099-4312, 4152-4403, 4158-4431, 4230-4401, 4230-4750, 4230-4811, 4250-4403, 4277-4817, 4299-4809, 4355-4590, 4355-4603, 4381-4784, 4460-4729, 4460-4816, 4460-4961, 4499-4777, 4501-4645, 4533-4776, 4633-4829, 4639-4874

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
75/1336022CB1/3298	1-494, 217-493, 220-494, 220-613, 220-630, 220-893, 223-831, 226-878, 228-833, 230-934, 234-874, 241-494, 244-613, 269-561, 360-613, 408-885, 465-658, 467-740, 495-1070, 549-1177, 566-827, 577-797, 633-895, 646-1194, 656-953, 668-857, 679-973, 694-885, 705-1428, 721-1177, 758-1079, 781-1362, 782-1009, 797-1044, 812-1122, 825-1101, 831-1122, 836-1108, 842-1070, 866-1100, 917-1125, 919-1610, 939-1177, 944-1561, 947-1222, 966-1234, 967-1134, 970-1436, 979-1535, 996-1232, 1017-1596, 1051-1629, 1074-1621, 1076-1623, 1076-1634, 1085-1345, 1097-1576, 1145-1637, 1159-1366, 1185-1436, 1200-1478, 1228-1671, 1239-1510, 1248-1399, 1249-1888, 1266-1573, 1355-1670, 1374-2001, 1375-1922, 1375-1976, 1424-1961, 1442-2078, 1468-2027, 1497-1977, 1543-2124, 1550-2186, 1624-2160, 1631-2137, 1637-2265, 1661-2205, 1665-2248, 1669-1961, 1678-1954, 1691-1962, 1691-2303, 1699-2305, 1715-2272, 1728-2335, 1740-1962, 1777-2348, 1840-2314, 1856-2086, 1862-2108, 1884-2510, 1885-2515, 1886-2213, 1906-2450, 1910-2426, 1954-1987, 1961-2213, 2017-2460, 2017-2551, 2042-2197, 2048-2507, 2092-2714, 2101-2314, 2167-2678, 2303-2839, 2376-2839, 2479-2895, 2479-2898, 2479-3109, 2538-3144, 2550-3282, 2550-3298, 2588-2879, 2589-2879, 2600-2839, 2603-2839, 2730-3002, 2841-3212
76/7473674CB1/833	1-321, 1-328, 44-463, 44-483, 46-367, 53-519, 68-236, 68-490, 91-537, 114-477, 128-415, 144-290, 269-653, 279-648, 282-485, 282-536, 282-583, 282-609, 282-654, 287-490, 288-654, 289-644, 289-654, 293-653, 300-620, 301-489, 301-645, 301-710, 319-727, 322-496, 323-446, 323-456, 323-470, 323-482, 323-618, 323-628, 323-630, 323-635, 323-637, 323-639, 323-645, 323-646, 323-647, 323-652, 323-653, 323-663, 323-717, 323-730, 323-731, 323-786, 324-647, 326-479, 331-507, 340-647, 341-730, 353-497, 369-497, 395-547, 395-638, 424-647, 426-654, 482-823, 504-654, 567-826, 640-833, 687-823
77/7475846CB1/920	1-920, 199-552
78/7475860CB1/964	1-553, 25-281, 240-535, 243-610, 255-872, 255-873, 268-872, 283-507, 285-849, 289-882, 291-592, 292-578, 293-511, 298-581, 298-722, 300-548, 301-563, 302-594, 309-753, 309-758, 309-760, 314-697, 314-783, 315-580, 315-591, 315-933, 316-678, 316-756, 316-785, 317-518, 317-580, 322-575, 322-586, 324-606, 325-570, 328-579, 330-614, 336-617, 342-623, 346-649, 347-607, 371-898, 377-946, 433-703, 477-820, 484-733, 493-964, 495-804, 512-910, 526-811, 568-807, 639-880, 639-964, 684-832, 762-962, 818-964
79/7950941CB1/701	1-397, 9-701
80/7485334CB1/1742	1-724, 126-1078, 752-995, 752-1279, 896-1188, 1002-1231, 1002-1507, 1038-1619, 1167-1454, 1167-1615, 1167-1742, 1282-1632
81/7220001CB1/2295	1-606, 1-762, 40-275, 149-553, 379-1700, 574-897, 736-1024, 736-1036, 738-1357, 1052-1657, 1117-1681, 1155-1690, 1159-1690, 1193-1839, 1202-1831, 1203-1796, 1226-1771, 1254-1820, 1324-1947, 1324-1969, 1347-1928, 1363-2014, 1367-2014, 1400-1867, 1423-1933, 1425-2084, 1432-2004, 1441-1867, 1499-2176, 1519-1700, 1547-2059, 1599-2093, 1600-2208, 1602-1806, 1605-2295, 1606-2201, 1645-2223, 2257-2285, 2257-2288, 2257-2290, 2257-2291, 2257-2292, 2257-2294, 2257-2295, 2258-2295

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
82/5956275CB1/911	1-676, 60-818, 61-349, 348-911
83/346472CB1/1806	1-259, 1-478, 1-1803, 115-677, 243-606, 292-495, 296-895, 313-977, 325-935, 353-997, 355-851, 361-1061, 362-935, 387-1063, 392-1010, 395-1077, 397-1029, 397-1051, 442-1006, 449-948, 463-979, 479-1190, 534-1032, 534-1100, 534-1102, 539-1042, 541-1191, 577-1094, 602-790, 615-950, 617-1298, 657-1047, 669-1293, 670-1298, 705-1110, 714-1314, 739-1367, 745-1271, 748-1367, 767-1282, 767-1286, 780-1367, 821-1116, 823-1421, 830-1481, 843-1555, 846-1443, 849-1506, 852-1496, 865-1404, 878-1528, 886-1563, 889-1596, 920-1451, 920-1460, 920-1463, 925-1227, 926-1392, 934-1475, 939-1595, 939-1625, 940-1483, 943-1504, 958-1171, 958-1519, 962-1479, 966-1566, 969-1468, 973-1314, 978-1489, 985-1511, 988-1417, 988-1535, 995-1675, 1001-1391, 1013-1286, 1013-1581, 1020-1170, 1122-1765, 1122-1793, 1167-1793, 1172-1443, 1188-1637, 1208-1793, 1214-1806, 1284-1767, 1294-1703, 1408-1684
84/643526CB1/603	1-539, 1-603, 4-603, 44-276, 44-535
85/1483418CB1/1888	1-105, 1-168, 1-193, 1-280, 1-394, 1-564, 1-631, 51-522, 89-618, 125-632, 125-644, 125-657, 129-776, 209-717, 211-751, 258-874, 290-680, 297-823, 343-964, 359-1005, 387-561, 396-968, 415-823, 450-507, 460-1056, 470-711, 470-996, 518-1009, 568-1020, 568-1024, 597-764, 606-910, 606-1200, 611-1300, 640-1093, 640-1194, 656-1181, 669-1268, 678-1005, 684-1331, 689-954, 696-1219, 696-1298, 711-1284, 713-1138, 768-1047, 768-1058, 768-1304, 781-1314, 842-1366, 895-1434, 918-1477, 940-1467, 1035-1561, 1112-1664, 1131-1383, 1267-1725, 1307-1554, 1330-1529, 1334-1808, 1334-1888, 1340-1528, 1340-1842, 1366-1642, 1407-1679, 1411-1872, 1430-1888
86/2683477CB1/1576	1-254, 1-513, 1-531, 1-537, 1-557, 1-592, 1-617, 1-641, 1-643, 1-653, 1-692, 1-713, 1-818, 2-742, 16-717, 27-967, 361-1008, 408-1254, 425-1121, 467-1168, 478-1175, 560-1141, 572-1380, 574-1399, 595-1085, 629-1195, 655-1307, 699-1314, 754-1196, 773-1576, 787-1284, 809-1573, 879-1576, 971-1573, 985-1576, 1060-1576, 1075-1576
87/5580991CB1/415	1-256, 1-415
88/5605931CB1/762	1-705, 66-705, 72-748, 157-705, 166-705, 211-705, 271-762, 275-759, 358-708, 366-701, 454-705, 480-705
89/6975241CB1/654	1-654, 96-653
90/6988529CB1/505	1-464, 1-505
91/6996808CB1/841	1-514, 85-548, 85-745, 85-841



Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
92/7472689CB1/1367	1-253, 192-448, 192-769, 192-781, 192-847, 284-331, 387-647, 444-546, 444-553, 444-587, 444-671, 444-674, 444-678, 444-699, 444-731, 444-734, 444-794, 444-810, 445-508, 481-1050, 483-963, 489-977, 490-765, 522-731, 529-863, 530-730, 543-937, 568-688, 572-1125, 604-859, 614-830, 627-1222, 630-1120, 663-1011, 667-1152, 672-1222, 686-1339, 715-1200, 721-927, 721-1259, 729-1290, 735-994, 759-1286, 770-1349, 796-1359, 818-1344, 837-1367, 861-1111, 861-1275, 861-1286, 861-1355, 873-1367, 890-1024, 897-1246, 897-1286, 909-1286, 911-1126, 912-1367, 914-1186, 922-1367, 941-1286, 944-1286, 951-1361, 955-1208, 960-1340, 1001-1304, 1001-1332, 1021-1286, 1022-1286, 1028-1286, 1035-1286, 1039-1367, 1041-1351, 1042-1295, 1075-1259, 1084-1286, 1094-1367, 1096-1329, 1101-1286, 1117-1286, 1118-1286, 1144-1286, 1152-1361, 1168-1286, 1180-1286, 1180-1287, 1202-1286, 1239-1286
93/876751CB1/4595	1-594, 248-705, 249-910, 277-740, 289-741, 403-907, 417-988, 440-992, 453-1111, 481-1192, 498-992, 503-1111, 568-812, 568-1066, 599-1188, 642-1192, 665-1303, 675-1202, 997-1564, 1045-1569, 1093-1340, 1093-1658, 1130-1681, 1197-1579, 1330-1617, 1375-1653, 1502-2181, 1507-2168, 1538-2103, 1674-1918, 1675-2130, 1678-2164, 1850-2311, 1890-2119, 1964-2399, 2120-2369, 2120-2746, 2130-2171, 2330-2538, 2330-2804, 2335-2624, 2348-2623, 2356-2650, 2423-2666, 2423-2844, 2447-2637, 2483-3218, 2498-2683, 2501-2670, 2521-2834, 2551-2816, 2563-2757, 2587-3211, 2592-2939, 2597-2878, 2637-3216, 2668-2996, 2694-2939, 2696-2946, 2696-3294, 2725-2994, 2811-3050, 2817-3230, 2839-3111, 2840-3203, 2843-3166, 2892-3145, 2892-3207, 2934-3222, 2945-3099, 2946-3181, 2950-3259, 2966-3244, 2985-3138, 2985-3180, 2995-3223, 2995-3470, 3002-3226, 3021-3324, 3053-3284, 3071-3360, 3080-3367, 3208-3464, 3208-3471, 3208-3720, 3215-3470, 3220-3498, 3233-3522, 3233-3528, 3257-3570, 3294-3845, 3341-3593, 3341-3625, 3350-3606, 3350-3625, 3379-3570, 3390-3653, 3391-3632, 3431-3570, 3432-3734, 3482-3726, 3482-3746, 3492-3730, 3515-3863, 3572-3827, 3572-3829, 3574-3867, 3574-4072, 3593-3852, 3646-3827, 3661-3922, 3661-3931, 3684-3921, 3694-3916, 3711-3981, 3714-3991, 3717-3992, 3743-4032, 3770-4030, 3788-4036, 3789-3948, 3799-4049, 3817-4036, 3820-4053, 3833-4085, 3859-4069, 3863-4027, 3863-4136, 3863-4376, 3866-4106, 3867-4180, 3869-3930, 3870-4126, 3877-4152, 3909-4123, 3940-4566, 3960-4571, 4045-4320, 4088-4379, 4095-4366, 4143-4560, 4160-4390, 4188-4579, 4216-4595, 4246-4465, 4263-4493, 4269-4525, 4366-4595, 4390-4569, 4425-4568, 4428-4574

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
94/2512510CB1/4759	1-218, 124-348, 124-403, 124-474, 124-511, 124-523, 124-529, 124-547, 124-576, 124-595, 124-612, 124-615, 124-621, 124-628, 124-631, 124-645, 124-689, 124-698, 124-713, 124-718, 124-724, 124-731, 124-734, 124-777, 124-778, 124-813, 124-852, 125-709, 126-398, 132-583, 142-594, 142-609, 155-583, 181-663, 197-433, 242-482, 256-561, 338-865, 365-597, 390-656, 436-671, 436-839, 473-882, 475-869, 514-959, 563-806, 689-1241, 813-1483, 833-1469, 945-1339, 946-1492, 973-1329, 1036-1523, 1114-1501, 1139-1426, 1176-1430, 1185-1684, 1212-1883, 1215-1465, 1216-1818, 1250-1811, 1273-1508, 1293-1876, 1313-1929, 1416-1880, 1462-2059, 1494-1803, 1494-1823, 1496-2069, 1512-1691, 1529-2126, 1530-1811, 1565-2214, 1568-2155, 1747-2231, 1808-2326, 1838-2387, 1866-2157, 1883-2594, 1893-2513, 1893-2547, 1901-2071, 2043-2312, 2071-2487, 2100-2372, 2192-2493, 2248-2427, 2272-2528, 2286-2552, 2292-2592, 2432-2722, 2483-2717, 2500-2743, 2538-2831, 2617-2750, 2626-2899, 2653-3239, 2654-2858, 2656-2908, 2730-3003, 2732-2978, 2781-3038, 2781-3316, 2866-3098, 2876-3553, 2887-3045, 2985-3264, 3024-3273, 3039-3276, 3051-3285, 3055-3331, 3075-3281, 3080-3362, 3099-3362, 3141-3440, 3146-3360, 3152-3420, 3153-3427, 3153-3441, 3181-3442, 3241-3496, 3344-3577, 3400-3661, 3400-3665, 3409-3663, 3424-3689, 3431-3762, 3449-3713, 3452-3738, 3460-3729, 3473-3731, 3481-3745, 3502-3775, 3504-3718, 3527-3775, 3528-3752, 3546-3786, 3548-3777, 3570-3853, 3582-3801, 3582-3806, 3617-3813, 3617-3877, 3637-3991, 3669-3877, 3683-3876, 3690-3898, 3707-3947, 3707-3964, 3709-3921, 3738-3994, 3738-4303, 3750-3970, 3783-4044, 3816-4101, 3844-4107, 3943-4256, 3960-4179, 3973-4187, 3973-4190, 3973-4191, 3973-4209, 3973-4374, 3984-4240, 3991-4236, 4000-4271, 4007-4216, 4007-4443, 4057-4295, 4072-4334, 4072-4336, 4072-4605, 4072-4733, 4094-4732, 4100-4759, 4102-4369, 4144-4707, 4154-4466, 4155-4393, 4155-4735, 4155-4745, 4165-4747, 4170-4747, 4171-4759, 4179-4444, 4184-4419, 4189-4759, 4205-4759, 4211-4759, 4213-4759, 4239-4748, 4241-4759, 4250-4747, 4258-4740, 4259-4747, 4266-4759, 4272-4754, 4292-4755, 4299-4747, 4300-4749, 4300-4756, 4302-4747, 4304-4747, 4305-4751, 4308-4747, 4313-4751, 4319-4759, 4320-4759, 4321-4740, 4323-4759, 4327-4747, 4329-4749, 4330-4741, 4330-4747, 4334-4753, 4335-4749, 4338-4753, 4338-4759, 4339-4753, 4339-4759, 4340-4740, 4340-4747, 4341-4747, 4343-4747, 4344-4747, 4345-4756, 4350-4747, 4354-4747, 4356-4750, 4360-4759, 4362-4582, 4362-4682, 4362-4710, 4362-4734, 4362-4743, 4375-4751, 4376-4747, 4381-4753, 4388-4748, 4391-4747, 4391-4750, 4396-4738, 4397-4747, 4400-4747, 4401-4751, 4403-4747, 4423-4747, 4427-4748, 4428-4747, 4429-4741, 4435-4649, 4435-4737, 4435-4738, 4435-4745, 4440-4747, 4444-4754, 4451-4728, 4456-4747, 4456-4753, 4462-4748, 4498-4759, 4503-4748, 4528-4748, 4536-4747, 4538-4752, 4549-4759, 4551-4750, 4551-4759, 4568-4747, 4570-4748, 4576-4759, 4584-4759, 4587-4731, 4604-4747, 4610-4741, 4612-4747, 4622-4748, 4625-4747, 4625-4748, 4640-4747, 4642-4747, 4644-4748, 4650-4747, 4662-4741, 4662-4747, 4662-4759, 4678-4747, 4685-4759
95/7486326CB1/3203	1-870, 411-3203

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
96/1221545CB1/1681	1-580, 69-613, 360-596, 377-763, 545-982, 569-775, 569-803, 583-1202, 657-901, 756-982, 784-1371, 986-1219, 1102-1615, 1204-1681, 1291-1511, 1339-1555, 1400-1614, 1446-1681
97/124737CB1/1207	1-529, 49-529, 309-1007, 319-729, 394-790, 402-566, 402-735, 406-568, 490-753, 502-988, 649-1179, 664-927, 671-1207, 697-963, 710-1092, 740-1175, 744-1092, 749-1092, 755-1015, 755-1207, 762-1207, 769-1179, 817-1008, 838-1092, 880-1092
98/1510784CB1/1544	1-196, 1-805, 316-894, 323-929, 366-783, 374-900, 374-916, 419-681, 446-962, 446-1005, 446-1038, 497-974, 534-886, 553-1008, 594-972, 658-1080, 669-1072, 717-1215, 1094-1523, 1094-1544, 1216-1447
99/1901257CB1/1519	1-248, 1-1519, 4-384, 79-583, 108-761, 481-645, 491-1007, 501-763, 545-1240, 708-1008, 714-1511, 784-1261, 787-1480, 802-1519, 808-1519, 817-1519, 839-1261, 839-1519, 841-1519, 853-1266, 878-1519, 901-1517, 924-1519, 931-1261, 945-1519, 1005-1519, 1007-1266, 1031-1261, 1241-1519
100/2044370CB1/525	1-235, 1-470, 1-525
101/2820933CB1/1062	1-862, 3-870, 431-1006, 440-1034, 440-1062, 443-995, 454-1009, 542-888, 663-966
102/2902793CB1/2155	1-2155, 66-129, 66-159, 66-304, 66-596, 66-653, 66-674, 92-736, 318-859, 318-860, 340-700, 372-973, 379-882, 379-923, 396-973, 405-973, 477-972, 487-937, 487-996, 487-998, 487-1053, 487-1054, 498-1083, 519-1002, 610-1001, 620-1016, 679-1189, 712-1259, 720-1096, 723-1252, 736-1436, 747-1252, 750-1252, 768-1309, 791-1178, 795-1052, 834-1122, 859-1420, 871-1380, 871-1385, 872-1355, 875-1252, 965-1287, 1002-1535, 1018-1597, 1047-1585, 1063-1463, 1157-1779, 1158-1586, 1184-1622, 1219-1704, 1240-1530, 1253-1717, 1253-1726, 1330-1933, 1333-1925, 1342-1977, 1386-1886, 1389-1631, 1394-1944, 1404-1977, 1446-2090, 1458-2078, 1487-2105, 1524-2071, 1621-2062, 1844-2155, 1978-2083
103/7486536CB1/1777	1-239, 1-259, 1-562, 10-255, 13-284, 19-248, 19-325, 19-451, 26-322, 49-265, 49-453, 140-367, 141-383, 148-368, 172-518, 174-553, 176-278, 181-281, 182-430, 182-454, 185-649, 189-449, 196-432, 196-825, 209-328, 217-486, 532-782, 634-914, 634-920, 634-1115, 634-1223, 644-926, 730-1013, 742-1003, 742-1013, 742-1055, 742-1300, 836-1135, 908-1128, 908-1377, 950-1192, 984-1199, 1055-1297, 1110-1390, 1117-1738, 1149-1753, 1171-1754, 1177-1663, 1181-1732, 1222-1733, 1302-1533, 1302-1754, 1302-1777, 1310-1758, 1324-1764, 1327-1622, 1333-1568, 1348-1755, 1454-1756, 1469-1696, 1511-1769, 1587-1765, 1590-1777, 1603-1765

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
104/8137305CB1/2587	1-308, 1-590, 1-899, 2-529, 5-481, 5-706, 27-611, 29-786, 30-611, 40-710, 41-276, 63-603, 87-873, 88-735, 89-688, 175-965, 309-609, 410-925, 481-1028, 550-806, 642-912, 676-1032, 693-936, 693-1024, 813-1102, 875-1451, 1042-1605, 1069-1228, 1170-1692, 1185-1403, 1290-1512, 1317-1947, 1346-1907, 1346-1998, 1381-1968, 1384-2097, 1427-2080, 1431-1686, 1463-1947, 1471-1955, 1548-1845, 1548-1991, 1559-2178, 1560-2068, 1566-1851, 1585-2176, 1592-1909, 1592-2009, 1604-1834, 1611-2217, 1674-2067, 1683-2151, 1689-2119, 1725-2276, 1734-1978, 1734-2119, 1734-2257, 1734-2301, 1734-2307, 1734-2587, 1740-2195, 1761-2201, 1775-2134, 1800-2211, 1802-2200
105/3793128CB1/1490	1-278, 1-452, 1-531, 1-534, 1-547, 1-583, 1-625, 1-707, 7-752, 91-728, 103-724, 110-524, 110-554, 110-575, 110-651, 110-677, 110-695, 110-739, 134-796, 193-576, 319-942, 435-1187, 436-970, 437-1067, 454-1190, 488-1163, 541-1134, 595-1270, 642-1173, 647-1162, 661-1280, 668-1271, 685-1271, 703-1376, 743-1435, 753-1286, 753-1454, 840-1490, 887-1349, 915-1374, 959-1484, 962-1484, 968-1437, 1005-1484, 1057-1484, 1073-1484, 1149-1484, 1177-1484, 1332-1484
106/4001243CB1/1174	1-292, 1-406, 1-442, 1-503, 1-570, 1-648, 52-303, 52-723, 86-644, 101-373, 101-583, 101-672, 113-321, 116-663, 285-854, 291-744, 291-898, 305-810, 415-1174, 427-710, 430-945, 436-723, 488-963, 504-720, 695-963
107/6986717CB1/818	1-556, 1-611, 13-546, 83-818
108/7503512CB1/4717	1-4679, 1-4690, 124-474, 124-511, 124-523, 124-529, 124-547, 124-576, 124-595, 124-599, 126-398, 132-583, 132-599, 142-594, 142-599, 155-583, 171-360, 197-433, 242-482, 252-357, 252-360, 257-561, 366-478, 396-946, 397-946, 620-1172, 628-1180, 744-1413, 764-1400, 848-1500, 876-1270, 877-1423, 904-1260, 967-1454, 1032-1696, 1038-1658, 1045-1432, 1079-1357, 1107-1361, 1116-1615, 1146-1396, 1147-1749, 1148-1397, 1152-1814, 1181-1742, 1197-1745, 1204-1439, 1224-1806, 1347-1811, 1399-1968, 1418-1492, 1425-1950, 1425-1977, 1427-2000, 1443-1622, 1461-1742, 1496-2145, 1678-2162, 1769-2318, 1809-2230, 2031-2303, 2773-3462, 2797-3170, 2799-3170, 2807-3484, 2820-3351, 2824-3468, 2935-3513, 2955-3204, 2968-3205, 2982-3216, 2986-3262, 2999-3194, 3008-3212, 3011-3293, 3019-3399, 3020-3261, 3022-3553, 3045-3499, 3057-3500, 3059-3500, 3072-3371, 3077-3291, 3084-3494, 3086-3372, 3093-3572, 3096-3532, 3110-3434, 3112-3373, 3120-3468, 3139-3605, 3152-3217, 3166-3290, 3216-3525, 3252-3499, 3276-3508, 3276-3880, 3279-3861, 3326-3572, 3331-3592, 3331-3596, 3340-3536, 3340-3594, 3340-3601, 3355-3620, 3357-3609, 3362-3693, 3379-4037, 3382-4151, 3383-3644, 3383-3669, 3384-3658, 3391-3660, 3396-3913, 3403-3651, 3404-3662,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
108 (cont.)	3404-3680, 3412-3676, 3433-3706, 3434-3728, 3434-3886, 3435-3649, 3443-3683, 3444-3742, 3456-3855, 3458-3694, 3458-3706, 3458-4047, 3459-3683, 3465-3750, 3477-3717, 3479-3708, 3491-4005, 3499-4117, 3501-3784, 3501-4125, 3513-3732, 3513-3737, 3515-3843, 3516-3806, 3528-3606, 3539-3830, 3542-3627, 3542-3807, 3542-3824, 3542-3882, 3542-4149, 3548-3744, 3548-3782, 3548-3808, 3568-3922, 3568-3978, 3571-4320, 3582-4060, 3586-3868, 3586-4184, 3588-4068, 3590-4295, 3600-3808, 3600-4261, 3603-3884, 3603-4075, 3603-4133, 3614-3807, 3621-3829, 3637-4192, 3638-3878, 3638-3895, 3640-3852, 3650-3933, 3654-4176, 3656-4212, 3659-3909, 3660-3950, 3667-3940, 3669-3925, 3669-4234, 3671-4053, 3671-4065, 3681-3901, 3707-4013, 3713-4166, 3714-3975, 3715-3971, 3717-3974, 3720-3873, 3721-3965, 3737-4296, 3738-4125, 3747-4032, 3756-4373, 3770-4118, 3772-4050, 3773-4212, 3775-4038, 3802-4095, 3808-4032, 3809-4390, 3813-4315, 3815-4059, 3815-4496, 3825-4084, 3825-4119, 3825-4120, 3828-4318, 3828-4462, 3829-4016, 3829-4043, 3831-4101, 3832-4343, 3838-4100, 3841-4091, 3847-4480, 3858-4082, 3860-4236, 3860-4303, 3866-4134, 3871-4132, 3874-4634, 3876-4187, 3877-4132, 3878-4164, 3892-4110, 3896-4608, 3901-4188, 3904-4118, 3904-4121, 3904-4122, 3904-4140, 3904-4305, 3904-4320, 3907-4157, 3908-4200, 3915-4171, 3920-4122, 3920-4345, 3922-4167, 3930-4225, 3931-4147, 3931-4186, 3931-4195, 3931-4201, 3931-4202, 3933-4165, 3934-4025, 3938-4147, 3938-4374, 3940-4201, 3942-4245, 3944-4238, 3947-4218, 3953-4561, 3956-4331, 3963-4628, 3967-4223, 3981-4265, 3982-4247, 3988-4226, 3988-4635, 3989-4545, 3990-4247, 3998-4211, 3998-4298, 3998-4673, 4000-4304, 4000-4606, 4003-4265, 4003-4267, 4003-4285, 4003-4536, 4003-4664, 4007-4230, 4013-4306, 4019-4552, 4025-4587, 4025-4631, 4025-4663, 4027-4289, 4031-4288, 4031-4310, 4031-4676, 4031-4705, 4033-4248, 4033-4253, 4033-4300, 4035-4561, 4043-4462, 4046-4341, 4049-4325, 4050-4280, 4052-4532, 4053-4324, 4059-4312, 4059-4356, 4071-4355, 4075-4638, 4084-4357, 4085-4642, 4086-4324, 4086-4676, 4087-4397, 4091-4329, 4094-4386, 4094-4663, 4094-4703, 4095-4346, 4096-4678, 4101-4678, 4103-4624, 4105-4574, 4109-4647, 4110-4375, 4111-4389, 4115-4350, 4115-4387, 4115-4398, 4115-4417, 4117-4359, 4120-4339, 4120-4674, 4128-4417, 4129-4425, 4133-4423, 4135-4367, 4138-4695, 4142-4697, 4144-4693, 4152-4486, 4170-4679, 4172-4711, 4174-4376, 4181-4678, 4189-4671, 4190-4459, 4190-4665, 4190-4678,

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
108 (cont.)	4190-4687, 4191-4363, 4192-4458, 4193-4439, 4197-4630, 4197-4702, 4200-4636, 4203-4685, 4206-4672, 4207-4678, 4212-4490, 4223-4686, 4230-4678, 4231-4676, 4231-4680, 4233-4678, 4235-4678, 4236-4682, 4239-4678, 4243-4619, 4244-4682, 4246-4586, 4250-4678, 4250-4692, 4251-4696, 4252-4671, 4254-4717, 4258-4678, 4260-4680, 4261-4672, 4261-4678, 4265-4684, 4266-4680, 4269-4684, 4269-4703, 4270-4684, 4270-4707, 4271-4671, 4271-4678, 4272-4678, 4274-4678, 4275-4678, 4276-4687, 4281-4678, 4285-4678, 4287-4681, 4289-4529, 4293-4513, 4293-4613, 4293-4641, 4293-4665, 4293-4674, 4295-4541, 4298-4581, 4306-4682, 4307-4678, 4312-4684, 4315-4558, 4319-4679, 4322-4678, 4322-4681, 4323-4597, 4327-4669, 4328-4678, 4331-4678, 4332-4682, 4334-4678, 4344-4597, 4351-4672, 4354-4678, 4355-4581, 4356-4625, 4358-4679, 4359-4678, 4360-4672, 4366-4580, 4366-4668, 4366-4669, 4366-4676, 4369-4655, 4375-4685, 4377-4478, 4377-4663, 4382-4659, 4387-4677, 4389-4668, 4393-4679, 4407-4644, 4429-4705, 4434-4679, 4459-4679, 4467-4678, 4469-4683, 4480-4708, 4482-4643, 4482-4681, 4482-4709, 4486-4575, 4499-4678, 4501-4679, 4515-4702, 4518-4662, 4535-4678, 4541-4672, 4556-4678, 4571-4678, 4573-4678, 4595-4672, 4595-4709, 4609-4678

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
55	095765CB1	PITUNOT06
56	6399886CB1	PANCTUT01
57	6024420CB1	TESTNOT11
58	7481067CB1	BRSTNOT07
59	3378720CB1	KERANOT02
60	938824CB1	CERVNOT01
61	1683721CB1	PROSNOT15
62	1694122CB1	COLNNOT23
63	1970615CB1	PROSTUT09
64	2314152CB1	CONUTUT01
65	2886225CB1	UTRSTMR02
66	6144418CB1	BRANDIN01
67	6834184CB1	BRSTNON02
68	6951005CB1	BRAITDR02
69	7250331CB1	KIDNTUT15
71	7011042CB1	BRAZNOT01
72	7427362CB1	BRSTTMR01
73	7485304CB1	SEMVTDE01
74	1422394CB1	MIXDUNB01
75	1336022CB1	COLNNOT13
76	7473674CB1	LUNGNON07
78	7475860CB1	ADRENON04
79	7950941CB1	BRABNOE02
80	7485334CB1	BSTMNON02
81	7220001CB1	COLXTDT01
82	5956275CB1	BRAUNOR01
83	346472CB1	THYMNOT02
84	643526CB1	BRAIFEE05
85	1483418CB1	SINTBST01
86	2683477CB1	SINIUCT01
87	5580991CB1	UTRENON03
88	5605931CB1	MONOTXN03
89	6975241CB1	BRAHTDR04
90	6988529CB1	BRAIFER05
91	6996808CB1	BRAXTDR17
92	7472689CB1	LNODNOT12
93	876751CB1	THYRNOT03
94	2512510CB1	BRAUNOR01
96	1221545CB1	NEUTFMT01
97	124737CB1	THYMNON04
98	1510784CB1	SINTFER03
99	1901257CB1	BRSTTMT02
100	2044370CB1	HIPONON02
101	2820933CB1	ADRETUT06
102	2902793CB1	DRGCNOT01
103	7486536CB1	BRSTNOT05
104	8137305CB1	MIXDTUE01

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
105	3793128CB1	BRSTNOT28
106	4001243CB1	PROSTMT03
107	6986717CB1	BRAIFER05
108	7503512CB1	BRSTNOT01



Table 6

Library	Vector	Library Description
ADRENON04	PSPORT1	Normalized library was constructed from 1.36 x 1e6 independent clones from an adrenal tissue library. Starting RNA was made from adrenal gland tissue removed from a 20-year-old Caucasian male, who died from head trauma. The library was normalized in two rounds using conditions adapted from Soares et al. (PNAS (1994) 91:9228-9232) and Bonaldo et al. (Genome Res (1996) 6: 791-806), using a significantly longer (48-hours/round) reannealing hybridization period.
ADRETUT06	pINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
BRABNOE02	PBK-CMV	This 5' biased random primed library was constructed using RNA isolated from vermis tissue removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly, and an enlarged spleen and liver. Patient medications included simethicone, Lasix, Digoxin, Colace, Zantac, captopril, and Vasotec.
BRAHTDR04	PCDNA2.1	This random primed library was constructed using RNA isolated archaocortex, anterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRAIFEE05	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
BRAIFER05	pINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
BRAITDR02	PCDNA2.1	This random primed library was constructed using RNA isolated from allocortex, neocortex, anterior and frontal cingulate tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.

Table 6 (cont.)

Library	Vector	Library Description
BRANDIN01	pINCY	<p>This normalized pineal gland tissue library was constructed from 4 million independent clones from a pineal gland tissue library from two different donors. Starting RNA was made from pooled pineal gland tissue removed from two Caucasian females: a 68-year-old (donor A) who died from congestive heart failure and a 79-year-old (donor B) who died from pneumonia.</p> <p>Neuropathology for donor A indicated mild to moderate Alzheimer disease, arteriolosclerosis, atherosclerosis, and multiple infarctions. Neuropathology for donor B indicated severe Alzheimer disease, arteriolosclerosis, cerebral amyloid angiopathy and multiple infarctions. There were diffuse and neuritic amyloid plaques and neurofibrillary tangles throughout the brain sections examined in both donors. Patient history included diabetes mellitus, rheumatoid arthritis, hyperthyroidism, amyloid heart disease, and dementia in donor A; and pseudophakia, gastritis with bleeding, glaucoma, peripheral vascular disease, COPD, delayed onset tonic/clonic seizures, and transient ischemic attack in donor B. The library was normalized in one round using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.</p>
BRAUNOR01	pINCY	<p>This random primed library was constructed using RNA isolated from striatum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloid goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease.</p>
BRAXTDRI7	PCDNA2.1	<p>This random primed library was constructed using RNA isolated from temporal neocortex tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.</p>

Table 6 (cont.)

Library	Vector	Library Description
BRAZNOT01	pINCY	Library was constructed using RNA isolated from striatum, globus pallidus and posterior putamen tissue removed from a 45-year-old Caucasian female who died from a dissecting aortic aneurysm and ischemic bowel disease. Pathology indicated mild arteriosclerosis involving the cerebral cortical white matter and basal ganglia. Grossly, there was mild meningeal fibrosis and mild focal atherosclerotic plaque in the middle cerebral artery, as well as vertebral arteries bilaterally. Microscopically, the cerebral hemispheres, brain stem and cerebellum reveal focal areas in the white matter that contain blood vessels that were barrel-shaped, hyalinized, with hemosiderin-laden macrophages in the Virchow-Robin space. In addition, there were scattered neurofibrillary tangles within the basolateral nuclei of the amygdala. Patient history included mild atheromatosis of aorta and coronary arteries, bowel and liver infarct due to aneurysm, physiologic fatty liver associated with obesity, mild diffuse emphysema, thrombosis of mesenteric and portal veins, cardiomegaly due to hypertrophy of left ventricle, arterial hypertension, acute pulmonary edema, splenomegaly, obesity (300 lb.), leiomyoma of uterus, sleep apnea, and iron deficiency anemia.
BRSTNON02	pINCY	This normalized breast tissue library was constructed from 6.2 million independent clones from a pool of two libraries from two different donors. Starting RNA was made from breast tissue removed from a 46-year-old Caucasian female during a bilateral reduction mammoplasty (donor A), and from breast tissue removed from a 60-year-old Caucasian female during a bilateral reduction mammoplasty (donor B). Pathology indicated normal breast parenchyma, bilaterally (A) and bilateral mammary hypertrophy (B). Patient history included hypertrophy of breast, obesity, lumbago, and glaucoma (A) and joint pain in the shoulder, thyroid cyst, colon cancer, normal delivery and cervical cancer (B). Family history included cataract, osteoarthritis, uterine cancer, benign hypertension, hyperlipidemia, and alcoholic cirrhosis of the liver, cerebrovascular disease, and type II diabetes (A) and cerebrovascular accident, atherosclerotic coronary artery disease, colon cancer, type II diabetes, hyperlipidemia, depressive disorder, and Alzheimer's Disease. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRSTNOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the breast tissue of a 56-year-old Caucasian female who died in a motor vehicle accident.
BRSTNOT05	PSPORT1	Library was constructed using RNA isolated from breast tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated multicentric invasive grade 4 lobular carcinoma. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular and cardiovascular disease, breast and prostate cancer, and type I diabetes.
BRSTNOT07	pINCY	Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.

Table 6 (cont.)

Library	Vector	Library Description
BRSTNOT28	pINCY	Library was constructed using RNA isolated from diseased right breast tissue removed from a 40-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated bilateral mild fibrocystic and proliferative changes. Patient history included pure hypercholesterolemia. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, type II diabetes, and prostate cancer.
BRSTTMR01	PCDNA2.1	This random primed library was constructed using RNA isolated from right breast tissue removed from a 62-year-old Caucasian female during open breast biopsy and unilateral extended simple mastectomy. Pathology indicated benign breast parenchyma. Pathology for the matched tumor tissue indicated residual grade 3 (of 4) ductal adenocarcinoma. The patient presented with breast cancer. Patient history included benign neoplasm of the large bowel and leg vein occlusion. Previous surgeries included dilation and curettage and spinal canal exploration. Patient medications included Lozal, Mevacor, and tamoxifen. Family history included heart murmur in the mother; skin cancer in the sibling(s); and prostate cancer in the grandfather.
BRSTTMT02	pINCY	Library was constructed using RNA isolated from diseased right breast tissue removed from a 46-year-old Caucasian female during a unilateral extended simple mastectomy and open breast biopsy. Pathology indicated mildly proliferative fibrocystic change, including intraductal duct ectasia, papilloma formation, and ductal hyperplasia. Pathology for the associated tumor tissue indicated multifocal ductal carcinoma in situ, both comedo and non-comedo types, nuclear grade 2 with extensive intraductal calcifications. Patient history included deficiency anemia, normal delivery, chronic sinusitis, extrinsic asthma, and kidney infection. Family history included type II diabetes, benign hypertension, cerebrovascular disease, skin cancer, and hyperlipidemia.
BSTMNON02	PSPORT1	This normalized brain stem library was constructed from 2.84 million independent clones from a brain stem library. Starting RNA was made from the brain stem tissue of a 72-year-old Caucasian male who died from myocardial infarction. Patient history included coronary artery disease, insulin-dependent diabetes mellitus, and arthritis. Normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228).
CERVNOT01	PSPORT1	Library was constructed using RNA isolated from the uterine cervical tissue of a 35-year-old Caucasian female during a vaginal hysterectomy with dilation and curettage. Pathology indicated mild chronic cervicitis. Family history included atherosclerotic coronary artery disease and type II diabetes.
COLNNOT13	pINCY	Library was constructed using RNA isolated from ascending colon tissue of a 28-year-old Caucasian male with moderate chronic ulcerative colitis.
COLNNOT23	pINCY	Library was constructed using RNA isolated from diseased colon tissue removed from a 16-year-old Caucasian male during a total colectomy with abdominal/perineal resection. Pathology indicated gastritis and pancolitis consistent with the acute phase of ulcerative colitis. Inflammation was more severe in the transverse colon, with inflammation confined to the mucosa. There was only mild involvement of the ascending and sigmoid colon, and no significant involvement of the cecum, rectum, or terminal ileum. Family history included irritable bowel syndrome.

Table 6 (cont.)

Library	Vector	Library Description
COLXTDT01	pINCY	Library was constructed using RNA isolated from colon tissue removed from the appendix of a 37-year-old Black female during myomectomy, dilation and curettage, right fimbrial region biopsy, and incidental appendectomy. Pathology indicated an unremarkable appendix. Pathology for the associated tumor tissue indicated multiple (12) uterine leiomyomata. Patient history included premenopausal menorrhagia and sarcoidosis of the lung. Family history included acute myocardial infarction and atherosclerotic coronary artery disease.
CONUTUT01	pINCY	Library was constructed using RNA isolated from sigmoid mesentery tumor tissue obtained from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed müllerian tumor present in the sigmoid mesentery at two sites.
DRGCNOT01	pINCY	Library was constructed using RNA isolated from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema and bronchopneumonia, bilateral pleural and pericardial effusions, and malignant lymphoma (natural killer cell type). Patient history included probable cytomegalovirus infection, hepatic congestion and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, and Bell's palsy. Surgeries included colonoscopy, large intestine biopsy, adenotonsillectomy, and nasopharyngeal endoscopy and biopsy; treatment included radiation therapy.
HIPONON02	PSPORT1	This normalized hippocampus library was constructed from 1.13M independent clones from a hippocampus tissue library. RNA was isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis. The normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228).
KERANOT02	PSPORT1	Library was constructed using RNA isolated from epidermal breast keratinocytes (NHEK). NHEK (Clontech #CC-2501) is human breast keratinocyte cell line derived from a 30-year-old black female during breast-reduction surgery.
KIDNTUT15	pINCY	Library was constructed using RNA isolated from left kidney tumor tissue removed from a 65-year-old Caucasian male during an exploratory laparotomy and nephroureterectomy. Pathology indicated grade 1 renal cell carcinoma within the upper pole of the left kidney. Patient history included malignant melanoma of the abdominal skin, benign neoplasm of colon, cerebrovascular disease, and umbilical hernia. Family history included myocardial infarction, atherosclerotic coronary artery disease, and cerebrovascular disease, and prostate cancer.
LNODNOT12	pINCY	Library was constructed using RNA isolated from lymph node tissue obtained from an 11-year-old Caucasian female who died from a motor vehicle accident. Previous surgeries included tonsilectomy.
LUNGNON07	pINCY	This normalized lung tissue library was constructed from 5.1 million independent clones from a lung tissue library. Starting RNA was made from RNA isolated from lung tissue. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

Table 6 (cont.)

Library	Vector	Library Description
MIXDTUE01	PBK-CMV	This 5' biased random primed library was constructed using pooled cDNA from seven donors. cDNA was generated using mRNA isolated from placental tissue removed from a Caucasian fetus (A), who died after 16 weeks' gestation from fetal demise and hydrocephalus; from placental tissue removed from a Caucasian male fetus (B), who died after 18 weeks' gestation from fetal demise; from an untreated LNCaP cell line, derived from prostate carcinoma with metastasis to the left supraclavicular lymph nodes, removed from a 50-year-old Caucasian male (C); from endometrial tissue removed from a 32-year-old female (D); from diseased right ovary tissue removed from a 45-year-old Caucasian female (E); from diseased right ovary tissue removed from a 47-year-old Caucasian female (donor F) and from right fallopian tube tumor tissue removed from an 85-year-old Caucasian female (donor G). For donor A, patient history included multiple pregnancies and live births, and an abortion in the mother. For donor B, serologies were negative. For donor D, pathology indicated the endometrium was in secretory phase. For donor E, pathology indicated stromal hyperthecosis of the right and left ovaries. For donor F, pathology indicated endometriosis. For donor G, pathology indicated poorly differentiated mixed endometrioid (80%) and serous (20%) adenocarcinoma of the right fallopian tube. Patient history included medullary carcinoma of the thyroid.
MIXDUNB01	pINCY	Library was constructed using RNA isolated from myometrium removed from a 41-year-old Caucasian female (A) during vaginal hysterectomy with a dilatation and curettage and untreated smooth muscle cells removed from the renal vein of a 57-year-old Caucasian male. Pathology for donor A indicated the myometrium and cervix were unremarkable. The endometrium was secretory and contained fragments of endometrial polyps. Benign endo- and ectocervical mucosa were identified in the endocervix. Pathology for the associated tumor tissue indicated uterine leiomyoma. Medical history included an unspecified menstrual disorder, ventral hernia, normal delivery, a benign ovarian neoplasm, and tobacco abuse in donor A. Previous surgeries included a bilateral destruction of fallopian tubes, removal of a solitary ovary, and an exploratory laparotomy in donor A. Medications included ferrous sulfate in donor A.
MONOTXN03	pINCY	Normalized, treated monocyte tissue library was constructed from 7.6 million independent clones from a treated monocyte library. Starting RNA was made from RNA isolated from treated monocytes from peripheral blood obtained from a 42-year-old female. The cells were treated with anti-interleukin-10 (anti-IL-10) and lipopolysaccharide (LPS). The anti-IL-10 was added at time 0 at 10 ng/ml and LPS was added at 1 hour at 5 ng/ml. The monocytes were isolated from buffy coat by adherence to plastic. Incubation time was 24 hours. cDNA synthesis was initiated using a NotI-anchored oligo(dT) primer. The libraries were normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48- hours/round) reannealing hybridization was used. The libraries were then linearized and recircularized to select for insert containing clones as follows: plasmid DNA was prepped from approximately 1 million clones from the normalized, treated monocyte tissue libraries following soft agar transformation. The DNA was linearized with NotI and insert containing clones were size-selected by agarose gel electrophoresis and recircularized by ligation.

Table 6 (cont.)

Library	Vector	Library Description
NEUTFMT01	PBLUESCRIPT	Library was constructed using total RNA isolated from peripheral blood granulocytes collected by density gradient centrifugation through Ficoll-Hypaque. The cells were isolated from buffy coat units obtained from unrelated male and female donors. Cells were cultured in 10 nM fMLP for 30 minutes, lysed in GuSCN, and spun through CsCl to obtain RNA for library construction. Because this library was made from total RNA, it has an unusually high proportion of unique singleton sequences, which may not all come from polyA RNA species.
PANCTUT01	pINCY	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 65-year-old Caucasian female during radical subtotal pancreatectomy. Pathology indicated an invasive grade 2 adenocarcinoma. Patient history included type II diabetes, osteoarthritis, cardiovascular disease, benign neoplasm in the large bowel, and a cataract. Previous surgeries included a total splenectomy, cholecystectomy, and abdominal hysterectomy. Family history included cardiovascular disease, type II diabetes, and stomach cancer.
PITUNOT06	pINCY	Library was constructed using RNA isolated from pituitary gland tissue removed from a 55-year-old male who died from chronic obstructive pulmonary disease. Neuropathology indicated there were no gross abnormalities, other than mild ventricular enlargement. There was no apparent microscopic abnormality in any of the neocortical areas examined, except for a number of silver positive neurons with apical dendrite staining, particularly in the frontal lobe. The significance of this was undetermined. The only other microscopic abnormality was that there was prominent silver staining with some swollen axons in the CA3 region of the anterior and posterior hippocampus. Microscopic sections of the cerebellum revealed mild Bergmann's gliosis in the Purkinje cell layer. Patient history included schizophrenia.
PROSNOT15	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 66-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 2+3). The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer, secondary bone cancer, and benign hypertension.
PROSTMT03	pINCY	The library was constructed using RNA isolated from right prostate tissue removed from a 68-year-old Caucasian male during a radical prostatectomy and regional lymph node excision. Pathology for the associated tumor indicated adenocarcinoma, Gleason grade 4+3, which formed a predominant mass involving the left side peripherally. The patient presented with elevated prostate specific antigen (PSA) and induration. Patient history included pure hypercholesterolemia, kidney calculus, an unspecified allergy, and atopidermatitis. Family history included colon cancer.
PROSTUT09	pINCY	Library was constructed using RNA isolated from prostate tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma. The patient presented with prostatic inflammatory disease. Patient history included lung neoplasm, and benign hypertension. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease and lung cancer.

Table 6 (cont.)

Library	Vector	Library Description
SEMYTDE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from seminal vesicle tissue removed from a 63-year-old Caucasian male during closed prostatic biopsy, radical prostatectomy, and regional lymph node excision. Pathology for the associated tumor tissue indicated Gleason grade 2+3 adenocarcinoma in the right side of the prostate. Adenofibromatous hyperplasia was present. The patient presented with prostate cancer, elevated prostate specific antigen and prostatic hyperplasia. Patient history included kidney calculus, extrinsic asthma, benign bowel neoplasm, backache, tremor, and tobacco abuse in remission. Previous surgeries included adenotomies. Patient medications included Ventolin and Vancril. Family history included atherosclerotic coronary artery disease and acute myocardial infarction in the mother; atherosclerotic coronary artery disease and acute myocardial infarction in the father; and stomach cancer and extrinsic asthma in the grandparent(s).
SINIUCT01	pINCY	Library was constructed using RNA isolated from ileum tissue obtained from a 42-year-old Caucasian male during a total intra-abdominal colectomy and endoscopic jejunostomy. Previous surgeries included polypectomy, colonoscopy, and spinal canal exploration. Family history included cerebrovascular disease, benign hypertension, atherosclerotic coronary artery disease, and type II diabetes.
SINTBST01	pINCY	Library was constructed using RNA isolated from ileum tissue obtained from an 18-year-old Caucasian female during bowel anastomosis. Pathology indicated Crohn's disease of the ileum, involving 15 cm of the small bowel. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
SINTFER03	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a Caucasian male fetus who died from fetal demise.
TESTNOT11	pINCY	Library was constructed using RNA isolated from testicular tissue removed from a 16-year-old Caucasian male who died from hanging. Patient history included drug use (tobacco, marijuana, and cocaine use), and medications included Lithium, Ritalin, and Paxil.
THYMNON04	PSPORT1	This normalized library was constructed from a thymus tissue library. Starting RNA was made from thymus tissue removed from a 3-year-old Caucasian male, who died from anoxia. Serologies were negative. The patient was not taking any medications. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48-hours/round) reannealing hybridization was used.
THYMNOT02	PBLUESCRIPT	Library was constructed using polyA RNA isolated from thymus tissue removed from a 3-year-old Caucasian male, who died from drowning. Serologies were negative.
THYRNOT03	pINCY	Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma, forming a well-encapsulated mass in the left thyroid.



Table 6 (cont.)

Library	Vector	Library Description
UTRENON03	pINCY	<p>This normalized library was constructed from <math>1.2 \times 10^6</math> independent clones from a uterine endometrial tissue library. Starting RNA was made from uterine endometrium tissue obtained from a 29-year-old Caucasian female during a vaginal hysterectomy and cystocele repair. Pathology indicated the endometrium was secretory and the cervix showed mild chronic cervicitis with focal squamous metaplasia. Pathology for the associated tumor tissue indicated an intramural uterine leiomyoma. Patient history included hypothyroidism, pelvic floor relaxation, incomplete T-12 injury (due to a motor vehicle accident) causing paraplegia and self catheterization. Previous surgeries included a normal delivery, a laminectomy, and a rhinoplasty. Family history included benign hypertension, type II diabetes, and hyperlipidemia. The libraries were normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.</p>
UTRSTMR02	PCDNA2.1	<p>This random primed library was constructed using pooled cDNA from two different donors. cDNA was generated using mRNA isolated from endometrial tissue removed from a 32-year-old female (donor A) and using mRNA isolated from myometrium removed from a 45-year-old female (donor B) during vaginal hysterectomy and bilateral salpingo-oophorectomy. In donor A, pathology indicated the endometrium was secretory phase. The cervix showed severe dysplasia (CIN III) focally involving the squamocolumnar junction at the 1, 6 and 7 o'clock positions. Mild koilocytotic dysplasia was also identified within the cervix. In donor B, pathology for the matched tumor tissue indicated multiple (23) subserosal, intramural, and submucosal leiomyomata. Patient history included stress incontinence, extrinsic asthma without status asthmaticus and normal delivery in donor B. Family history included cerebrovascular disease, depression, and atherosclerotic coronary artery disease in donor B.</p>

Table 7

Program	Description	Reference	Parameter Threshold
ABIFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score $\geq$ GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
  - a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54,
  - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-53,
  - c) a polypeptide consisting essentially of a naturally occurring amino acid sequence at least 91% identical to the amino acid sequence of SEQ ID NO:54,
  - d) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, and
  - e) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-54.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method of producing a polypeptide of claim 1, the method comprising:
  - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant

polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

5 10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-54.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

10 12. An isolated polynucleotide selected from the group consisting of:

- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:55-108,
- b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of
- 15 SEQ ID NO:55-108,
- c) a polynucleotide complementary to a polynucleotide of a),
- d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).

20 13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- 25 a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- 30 b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

35 16. A method of detecting a target polynucleotide in a sample, said target polynucleotide

having a sequence of a polynucleotide of claim 12, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-54.

19. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition of claim 17.

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

25. A method for treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 24.

26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions

whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

- c) quantifying the amount of hybridization complex, and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A diagnostic test for a condition or disease associated with the expression of SECP in a biological sample, the method comprising:

- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
- b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

31. The antibody of claim 11, wherein the antibody is:

- a) a chimeric antibody,
- b) a single chain antibody,
- c) a Fab fragment,
- d) a  $F(ab')_2$  fragment, or
- e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

33. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, wherein the antibody is labeled.

35. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 34.



36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibodies from said animal, and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54.

37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-54, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- d) culturing the hybridoma cells, and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant

immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54 in a sample, the method comprising:

- 5           a)       incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b)       detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54 in the sample.

10

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54 from a sample, the method comprising:

- a)       incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- 15           b)       separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-54.

20           46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

- 25           a)       labeling the polynucleotides of the sample,
- b)       contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
- c)       quantifying the expression of the polynucleotides in the sample.

30           48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

35

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is  
5 completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

10 52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

15 54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical  
20 location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

25 56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

30 59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

35 61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

5 64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

10

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

15

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

20

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

25

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

30

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

35

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

5 82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

10 84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

15 87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

20 89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

25 92. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

93. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

30 94. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

95. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

96. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

35 97. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

98. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.

99. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.

5 100. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:45.

101. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:46.

102. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:47.

10

103. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:48.

104. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:49.

15

105. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:50.

106. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:51.

107. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:52.

20

108. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:53.

109. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:54.

25

110. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:55.

111. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:56.

30

112. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:57.

35

113. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:58.

114. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:59.

115. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
5 NO:60.

116. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:61.

117. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
10 NO:62.

118. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:63.

119. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
15 NO:64.

120. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
20 NO:65.

121. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:66.

122. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
25 NO:67.

123. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:68.

124. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
30 NO:69.

125. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
35 NO:70.

126. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:71.

127. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
5 NO:72.

128. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:73.

129. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
10 NO:74.

130. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:75.

131. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
15 NO:76.

132. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
20 NO:77.

133. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:78.

134. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
25 NO:79.

135. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:80.

136. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
30 NO:81.

137. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
35 NO:82.



138. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:83.

139. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
5 NO:84.

140. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:85.

10 141. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:86.

142. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:87.

15 143. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:88.

20 144. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:89.

145. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:90.

25 146. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:91.

147. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:92.

30 148. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:93.

35 149. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:94.

150. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:95.

5 151. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:96.

152. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:97.

10 153. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:98.

154. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:99.

15 155. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:100.

20 156. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:101.

157. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:102.

25 158. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:103.

159. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:104.

30 160. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:105.

35 161. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
NO:106.

162. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:107.

163. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID  
5 NO:108.

<110> INCYTE GENOMICS, INC.  
 GRIFFIN, Jennifer A.  
 YAO, Monique G.  
 DUGGAN, Brendan M.  
 YUE, Henry  
 DING, Li  
 LAL, Preeti G.  
 LEE, Ernestine A.  
 RAMKUMAR, Jayalaxmi  
 THANGAVELU, Kavitha  
 XU, Yuming  
 LEE, Sally  
 TANG, Y. Tom  
 NGUYEN, Danniel B.  
 WARREN, Bridget A.  
 HONCHELL, Cynthia D.  
 GIETZEN, Kimberly J.  
 BAUGHN, Mariah R.  
 GANDHI, Ameena R.  
 ARVIZU, Chandra  
 WALIA, Narinder K.  
 LU, Yan  
 ELLIOTT, Vicki S.  
 LU, Dyung Aina M.  
 HAFALIA, April J.A.  
 AZIMZAI, Yalda  
 KHAN, Farrah A.  
 UYEN, K. Tran

<120> SECRETED PROTEINS

<130> PI-0345 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/255,639; 60/257,852; 60/260,105; 60/262,932; 60/263,096;  
 60/263,090; 60/265,926

<151> 2000-12-13; 2000-12-21; 2001-01-05; 2001-01-18; 2000-01-18;  
 2001-01-19; 2001-02-02

<160> 108

<170> PERL Program

<210> 1

<211> 235

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 095765CD1

<400> 1

Met	Pro	Arg	Ser	Cys	Cys	Ser	Arg	Ser	Gly	Ala	Leu	Leu	Leu	Ala
1				5					10					15
Leu	Leu	Leu	Gln	Ala	Ser	Met	Glu	Val	Arg	Gly	Trp	Cys	Leu	Glu
			20						25					30
Ser	Ser	Gln	Cys	Gln	Asp	Leu	Thr	Thr	Glu	Ser	Asn	Leu	Leu	Glu
			35						40					45
Cys	Ile	Arg	Ala	Cys	Lys	Pro	Asp	Leu	Ser	Ala	Glu	Thr	Pro	Met
			50						55					60

Phe Pro Gly Asn Gly Asp Glu Gln Pro Leu Thr Glu Asn Pro Arg  
 65 70 75  
 Lys Tyr Val Met Gly His Phe Arg Trp Asp Arg Phe Gly Arg Arg  
 80 85 90  
 Asn Ser Ser Asp Gly Ala Lys Pro Gly Pro Arg Glu Gly Lys Arg  
 95 100 105  
 Ser Tyr Ser Met Glu His Phe Arg Trp Gly Lys Pro Val Gly Lys  
 110 115 120  
 Lys Arg Arg Pro Val Lys Val Tyr Pro Asn Gly Ala Glu Asp Glu  
 125 130 135  
 Ser Ala Glu Ala Phe Pro Leu Glu Phe Lys Arg Glu Leu Thr Gly  
 140 145 150  
 Gln Arg Leu Arg Glu Gly Asp Gly Pro Asp Gly Pro Ala Asp Asp  
 155 160 165  
 Gly Ala Gly Ala Gln Ala Asp Leu Glu His Ser Leu Leu Val Ala  
 170 175 180  
 Ala Glu Lys Lys Asp Glu Gly Pro Tyr Arg Met Glu His Phe Arg  
 185 190 195  
 Trp Gly Ser Pro Pro Lys Asp Lys Arg Tyr Gly Gly Phe Met Thr  
 200 205 210  
 Ser Glu Lys Ser Gln Thr Pro Leu Val Thr Leu Phe Lys Asn Ala  
 215 220 225  
 Ile Ile Lys Asn Ala Tyr Lys Lys Gly Glu  
 230 235

<210> 2  
 <211> 689  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6399886CD1

<400> 2  
 Met Ala Ala Arg Thr Leu Gly Arg Gly Val Gly Arg Leu Leu Gly  
 1 5 10 15  
 Ser Leu Arg Gly Leu Ser Gly Gln Pro Ala Arg Pro Pro Cys Gly  
 20 25 30  
 Val Ser Ala Pro Arg Arg Ala Ala Ser Gly Pro Ser Gly Ser Ala  
 35 40 45  
 Pro Ala Val Ala Ala Ala Ala Ala Gln Pro Gly Ser Tyr Pro Ala  
 50 55 60  
 Leu Ser Ala Gln Ala Ala Arg Glu Pro Ala Ala Phe Trp Gly Pro  
 65 70 75  
 Leu Ala Arg Asp Thr Leu Val Trp Asp Thr Pro Tyr His Thr Val  
 80 85 90  
 Trp Asp Cys Asp Phe Ser Thr Gly Lys Ile Gly Trp Phe Leu Gly  
 95 100 105  
 Gly Gln Leu Asn Val Ser Val Asn Cys Leu Asp Gln His Val Arg  
 110 115 120  
 Lys Ser Pro Glu Ser Val Ala Leu Ile Trp Glu Arg Asp Glu Pro  
 125 130 135  
 Gly Thr Glu Val Arg Ile Thr Tyr Arg Glu Leu Leu Glu Thr Thr  
 140 145 150  
 Cys Arg Leu Ala Asn Thr Leu Lys Arg His Gly Val His Arg Gly  
 155 160 165  
 Asp Arg Val Ala Ile Tyr Met Pro Val Ser Pro Leu Ala Val Ala  
 170 175 180  
 Ala Met Leu Ala Cys Ala Arg Ile Gly Ala Val His Thr Val Ile  
 185 190 195  
 Phe Ala Gly Phe Ser Ala Glu Ser Leu Ala Gly Arg Ile Asn Asp  
 200 205 210

Ala Lys Cys Lys	Val Val Ile Thr Phe	Asn Gln Gly Leu Arg Gly	215	220	225
Gly Arg Val Val	Glu Leu Lys Lys Ile	Val Asp Glu Ala Val Lys	230	235	240
His Cys Pro Thr	Val Gln His Val Leu	Val Ala His Arg Thr Asp	245	250	255
Asn Lys Val His	Met Gly Asp Leu Asp	Val Pro Leu Glu Gln Glu	260	265	270
Met Ala Lys Glu	Asp Pro Val Cys Ala	Pro Glu Ser Met Gly Ser	275	280	285
Glu Asp Met Leu	Phe Met Leu Tyr Thr	Ser Gly Ser Thr Gly Met	290	295	300
Pro Lys Gly Ile	Val His Thr Gln Ala	Gly Tyr Leu Leu Tyr Ala	305	310	315
Ala Leu Thr His	Lys Leu Val Phe Asp	His Gln Pro Gly Asp Ile	320	325	330
Phe Gly Cys Val	Ala Asp Ile Gly Trp	Ile Thr Gly His Ser Tyr	335	340	345
Val Val Tyr Gly	Pro Leu Cys Asn Gly	Ala Thr Ser Val Leu Phe	350	355	360
Glu Ser Thr Pro	Val Tyr Pro Asn Ala	Gly Arg Tyr Trp Glu Thr	365	370	375
Val Glu Arg Leu	Lys Ile Asn Gln Phe	Tyr Gly Ala Pro Thr Ala	380	385	390
Val Arg Leu Leu	Leu Lys Tyr Gly Asp	Ala Trp Val Lys Lys Tyr	395	400	405
Asp Arg Ser Ser	Leu Arg Thr Leu Gly	Ser Val Gly Glu Pro Ile	410	415	420
Asn Cys Glu Ala	Trp Glu Trp Leu His	Arg Val Val Gly Asp Ser	425	430	435
Arg Cys Thr Leu	Val Asp Thr Trp Trp	Gln Thr Glu Thr Gly Gly	440	445	450
Ile Cys Ile Ala	Pro Arg Pro Ser Glu	Glu Gly Ala Glu Ile Leu	455	460	465
Pro Ala Met Ala	Met Arg Pro Phe Phe	Gly Ile Val Pro Val Leu	470	475	480
Met Asp Glu Lys	Gly Ser Val Met Glu	Gly Ser Asn Val Ser Gly	485	490	495
Ala Leu Cys Ile	Ser Gln Ala Trp Pro	Gly Met Ala Arg Thr Ile	500	505	510
Tyr Gly Asp His	Gln Arg Phe Val Asp	Ala Tyr Phe Lys Ala Tyr	515	520	525
Pro Gly Tyr Tyr	Phe Thr Gly Asp Gly	Ala Tyr Arg Thr Glu Gly	530	535	540
Gly Tyr Tyr Gln	Ile Thr Gly Arg Met	Asp Asp Val Ile Asn Ile	545	550	555
Ser Gly His Arg	Leu Gly Thr Ala Glu	Ile Glu Asp Ala Ile Ala	560	565	570
Asp His Pro Ala	Val Pro Glu Ser Ala	Val Ile Gly Tyr Pro His	575	580	585
Asp Ile Lys Gly	Glu Ala Ala Phe Ala	Phe Ile Val Val Lys Asp	590	595	600
Ser Ala Gly Asp	Ser Asp Val Val Val	Gln Glu Leu Lys Ser Met	605	610	615
Val Ala Thr Lys	Ile Ala Lys Tyr Ala	Val Pro Asp Glu Ile Leu	620	625	630
Val Val Lys Arg	Leu Pro Lys Thr Arg	Ser Gly Lys Val Met Arg	635	640	645
Arg Leu Leu Arg	Lys Ile Ile Thr Ser	Glu Ala Gln Glu Leu Gly	650	655	660
Asp Thr Thr Thr	Leu Glu Asp Pro Ser	Ile Ile Ala Glu Ile Leu	665	670	675
Ser Val Tyr Gln	Lys Cys Lys Asp Lys	Gln Ala Ala Ala Lys			

680

685

<210> 3  
 <211> 584  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID.No: 6024420CD1

<400> 3  
 Met Asp Leu Leu Trp Met Pro Leu Leu Leu Val Ala Ala Cys Val  
 1 5 10 15  
 Ser Ala Val His Ser Ser Pro Glu Val Asn Ala Gly Val Ser Ser  
 20 25 30  
 Ile His Ile Thr Lys Pro Val His Ile Leu Glu Glu Arg Ser Leu  
 35 40 45  
 Leu Val Leu Thr Pro Ala Gly Leu Thr Gln Met Leu Asn Gln Thr  
 50 55 60  
 Arg Phe Leu Met Val Leu Phe His Asn Pro Ser Ser Lys Gln Ser  
 65 70 75  
 Arg Asn Leu Ala Glu Glu Leu Gly Lys Ala Val Glu Ile Met Gly  
 80 85 90  
 Lys Gly Lys Asn Gly Ile Gly Phe Gly Lys Val Asp Ile Thr Ile  
 95 100 105  
 Glu Lys Glu Leu Gln Gln Glu Phe Gly Ile Thr Lys Ala Pro Glu  
 110 115 120  
 Leu Ser Cys Phe Leu Arg Ala Thr Arg Ser Glu Pro Ile Ser Cys  
 125 130 135  
 Lys Gly Val Val Glu Ser Ala Ala Leu Val Val Trp Leu Arg Arg  
 140 145 150  
 Gln Ile Ser Gln Lys Ala Phe Leu Phe Asn Ser Ser Glu Gln Val  
 155 160 165  
 Ala Glu Phe Val Ile Ser Arg Pro Leu Val Ile Val Gly Phe Phe  
 170 175 180  
 Gln Asp Leu Glu Glu Glu Val Ala Glu Leu Phe Tyr Asp Val Ile  
 185 190 195  
 Lys Asp Phe Pro Glu Leu Thr Phe Gly Val Ile Thr Ile Gly Asn  
 200 205 210  
 Val Ile Gly Arg Phe His Val Thr Leu Asp Ser Val Leu Val Phe  
 215 220 225  
 Lys Lys Gly Lys Ile Val Asn Arg Gln Lys Leu Ile Asn Asp Ser  
 230 235 240  
 Thr Asn Lys Gln Glu Leu Asn Arg Val Ile Lys Gln His Leu Thr  
 245 250 255  
 Asp Phe Val Ile Glu Tyr Asn Thr Glu Asn Lys Asp Leu Ile Ser  
 260 265 270  
 Glu Leu His Ile Met Ser His Met Leu Leu Phe Val Ser Lys Ser  
 275 280 285  
 Ser Glu Ser Tyr Gly Ile Ile Ile Gln His Tyr Lys Leu Ala Ser  
 290 295 300  
 Lys Glu Phe Gln Asn Lys Ile Leu Phe Ile Leu Val Asp Ala Asp  
 305 310 315  
 Glu Pro Arg Asn Gly Arg Val Phe Lys Tyr Phe Arg Val Thr Glu  
 320 325 330  
 Val Asp Ile Pro Ser Val Gln Ile Leu Asn Leu Ser Ser Asp Ala  
 335 340 345  
 Arg Tyr Lys Met Pro Ser Asp Asp Ile Thr Tyr Glu Ser Leu Lys  
 350 355 360  
 Lys Phe Gly Arg Ser Phe Leu Ser Lys Asn Ala Thr Lys His Gln  
 365 370 375  
 Ser Ser Glu Glu Ile Pro Lys Tyr Trp Asp Gln Gly Leu Val Lys-

	380		385		390
Gln Leu Val Gly	Lys Asn Phe Asn Val	Val Val Phe Asp Lys	Glu		
	395		400		405
Lys Asp Val Phe	Val Met Phe Tyr Ala	Pro Trp Ser Lys Lys	Cys		
	410		415		420
Lys Met Leu Phe	Pro Leu Leu Glu Glu	Leu Gly Arg Lys Tyr	Gln		
	425		430		435
Asn His Ser Thr	Ile Ile Ile Ala Lys	Ile Asp Val Thr Ala	Asn		
	440		445		450
Asp Ile Gln Leu	Met Tyr Leu Asp Arg	Tyr Pro Phe Phe Arg	Leu		
	455		460		465
Phe Pro Ser Gly	Ser Gln Gln Ala Val	Leu Tyr Lys Gly Glu	His		
	470		475		480
Thr Leu Lys Gly	Phe Ser Asp Phe Leu	Glu Ser His Ile Lys	Thr		
	485		490		495
Lys Ile Glu Asp	Glu Asp Glu Leu Leu	Ser Val Glu Gln Asn	Glu		
	500		505		510
Val Ile Glu Glu	Glu Val Leu Ala Glu	Glu Lys Glu Val Pro	Met		
	515		520		525
Met Lys Lys Glu	Leu Pro Glu Gln Gln	Ser Pro Glu Leu Glu	Asn		
	530		535		540
Met Thr Lys Tyr	Val Ser Lys Leu Glu	Glu Pro Ala Gly Lys	Lys		
	545		550		555
Lys Thr Ser Glu	Glu Val Val Val Val	Val Ala Lys Pro Lys	Gly		
	560		565		570
Pro Pro Val Gln	Lys Lys Lys Pro Lys	Val Lys Glu Glu Leu			
	575		580		

&lt;210&gt; 4

&lt;211&gt; 1049

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7481067CD1

&lt;400&gt; 4

Met Lys Ala Leu Leu	Pro Leu Thr Phe Leu	Phe Phe Ile Ser Ser	
1	5	10	15
Pro Gly Trp Ala Ile	Asp Arg His Cys Tyr	Ile Gly Ile Glu Glu	
	20	25	30
Ser Ile Trp Asn Tyr	Ala Pro Ser Gly Lys	Asn Met Leu Asn Glu	
	35	40	45
Lys Pro Phe Ser Glu	Asp Leu Glu Phe Leu	Gln Gly Gly Gln Ala	
	50	55	60
Arg Lys Ser Phe Val	Phe Lys Lys Ala Leu	Tyr Phe Gln Tyr Thr	
	65	70	75
Asp Asn Thr Phe Gln	Arg Ile Ile Glu Lys	Pro Ser Trp Leu Gly	
	80	85	90
Phe Leu Gly Pro Met	Ile Lys Ala Glu Thr	Gly Asp Phe Ile Tyr	
	95	100	105
Val His Val Lys Asn	Asn Ala Ser Arg Ala	Tyr Ser Tyr His Pro	
	110	115	120
His Gly Leu Thr Tyr	Ser Lys Glu Asn Glu	Gly Ala Ile Tyr Pro	
	125	130	135
Asp Asn Thr Thr Gly	Leu Gln Lys Glu Asp	Glu Tyr Leu Glu Pro	
	140	145	150
Gly Lys Gln Tyr Thr	Tyr Lys Trp Tyr Val	Glu Glu His Gln Gly	
	155	160	165
Pro Gly Pro Asn Asp	Ser Asn Cys Val Thr	Arg Ile Tyr His Ser	
	170	175	180
His Ile Asp Thr Ala	Arg Asp Val Ala Ser	Gly Leu Ile Gly Pro	



Ile Leu Thr Cys	185	190	195
Lys Arg Gly Thr Leu	200	205	210
Asp Ile Asp Arg	215	220	225
Ser Arg Ser Trp	230	235	240
Ser Gly Lys Ile	245	250	255
Ser Met Gln Ala	260	265	270
Leu Thr Met Cys	275	280	285
Met Gly Gly Val	290	295	300
Thr Leu Ile Ser	305	310	315
Pro Ser Ser Leu	320	325	330
Val Trp Met Leu	335	340	345
Phe Phe Lys Val	350	355	360
Val Thr Gly Thr	365	370	375
Ile Leu Trp Asn	380	385	390
Lys Asn Leu Thr	395	400	405
Arg Ser Pro Thr	410	415	420
Arg Glu Tyr Thr	425	430	435
Glu His Leu Gly	440	445	450
Gln Thr Ile Lys	455	460	465
Ser Ile Gln Pro	470	475	480
Leu Phe Tyr Glu	485	490	495
Val Ser Pro Gly	500	505	510
Asp Val Gly Pro	515	520	525
Tyr Tyr Ser Ser	530	535	540
Leu Gly Pro Leu	545	550	555
Gly Lys Gln Lys	560	565	570
Ile Phe Asp Glu	575	580	585
Thr Phe Ile Thr	590	595	600
Cys Gln Ala Ser	605	610	615
Gly Asn Leu Pro	620	625	630
Trp His Val Phe	635	640	645
Tyr Phe Ser Gly	650	655	660

Thr	Ile	Pro	Met	Phe	Pro	Tyr	Thr	Ser	Gln	Thr	Leu	Leu	Met	Thr
				665					670					675
Pro	Asp	Ser	Ile	Gly	Thr	Phe	Asp	Leu	Val	Cys	Met	Thr	Ile	Lys
				680					685					690
His	Asn	Leu	Gly	Gly	Met	Lys	His	Lys	Tyr	His	Val	Arg	Gln	Cys
				695					700					705
Gly	Lys	Pro	Asn	Pro	Asp	Gln	Thr	Gln	Tyr	Gln	Glu	Glu	Lys	Ile
				710					715					720
Ile	Ile	Thr	Ile	Ala	Ala	Glu	Glu	Met	Glu	Trp	Asp	Tyr	Ser	Pro
				725					730					735
Ser	Arg	Lys	Trp	Glu	Asn	Glu	Leu	His	His	Leu	Arg	Arg	Glu	Gln
				740					745					750
Thr	Ser	Met	Tyr	Val	Asp	Arg	Ser	Gly	Thr	Leu	Leu	Gly	Ser	Lys
				755					760					765
Tyr	Lys	Lys	Val	Leu	Tyr	Arg	Gln	Tyr	Asp	Asp	Asn	Thr	Phe	Thr
				770					775					780
Asn	Gln	Thr	Lys	Arg	Asn	Glu	Gly	Glu	Lys	His	Leu	Asp	Ile	Leu
				785					790					795
Gly	Pro	Leu	Ile	Leu	Leu	Asn	Pro	Gly	Gln	Ile	Ile	Gln	Ile	Ile
				800					805					810
Phe	Lys	Asn	Lys	Ala	Ala	Arg	Pro	Tyr	Ser	Ile	His	Ala	His	Gly
				815					820					825
Val	Lys	Thr	Asn	Asn	Ser	Thr	Val	Val	Pro	Thr	Gln	Pro	Gly	Glu
				830					835					840
Ile	Gln	Ile	Tyr	Thr	Trp	Gln	Ile	Pro	Asp	Arg	Thr	Gly	Pro	Thr
				845					850					855
Ser	Leu	Asp	Phe	Glu	Cys	Ile	Pro	Trp	Phe	Tyr	Tyr	Ser	Thr	Val
				860					865					870
Ser	Val	Ala	Lys	Asp	Leu	His	Ser	Gly	Leu	Val	Gly	Pro	Leu	Ser
				875					880					885
Val	Cys	Arg	Lys	Asp	Ile	Asn	Pro	Asn	Ile	Val	His	Arg	Val	Leu
				890					895					900
His	Phe	Met	Ile	Phe	Asp	Glu	Asn	Glu	Ser	Trp	Tyr	Phe	Glu	Asp
				905					910					915
Ser	Ile	Asn	Thr	Tyr	Ala	Ser	Lys	Pro	Asn	Lys	Val	Asp	Lys	Glu
				920					925					930
Asn	Asp	Asn	Phe	Gln	Leu	Ser	Asn	Gln	Met	His	Ala	Ile	Asn	Gly
				935					940					945
Arg	Leu	Phe	Gly	Asn	Asn	Gln	Gly	Ile	Thr	Phe	His	Val	Gly	Asp
				950					955					960
Val	Val	Asn	Trp	Tyr	Leu	Ile	Gly	Ile	Gly	Asn	Glu	Ala	Asp	Leu
				965					970					975
His	Thr	Val	His	Phe	His	Gly	His	Ser	Phe	Glu	Tyr	Lys	Asn	Arg
				980					985					990
Gly	Val	Tyr	Gln	Ser	Asp	Val	Tyr	Asp	Leu	Pro	Pro	Gly	Val	Tyr
				995					1000					1005
Arg	Thr	Val	Lys	Met	Tyr	Arg	Arg	Asp	Val	Gly	Thr	Trp	Leu	Phe
				1010					1015					1020
Tyr	Cys	His	Val	Phe	Glu	His	Ile	Gly	Ala	Gly	Met	Glu	Ser	Thr
				1025					1030					1035
Tyr	Thr	Val	Leu	Glu	Arg	Lys	Gly	Leu	Met	Glu	Gln	Asn	Leu	
				1040					1045					

&lt;210&gt; 5

&lt;211&gt; 383

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3378720CD1

&lt;400&gt; 5

```
<210> 6
<211> 72
<212> PRT
<213> Homo sapiens
```

8175

```

Met Pro Ala Ser Leu Trp Ala Phe Pro Arg Lys Lys His Trp Phe
 1           5           10           15
Leu Ser Ile Val Pro Trp Leu Val Leu Phe Leu Thr Leu Gly Leu
           20           25           30
Cys Val Arg Asn Lys Ala Ala Lys Leu His Val Val Ile Gln Gln
           35           40           45
Lys Glu Tyr Ser Asp Leu Ser Phe Ile Leu Leu Ile Val Pro Ser
           50           55           60
Thr Pro Ala Ala Ala Pro Ala Lys Tyr Tyr His Pro
           65           70

```

<210> 7  
 <211> 91  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1683721CD1

```

<400> 7
Met Met Leu Gly Trp Gly Trp Lys Ala Leu Leu Leu Lys Ser Leu
 1           5           10           15
Ala Phe Pro Thr Gln Gly Tyr Pro Glu Gly Tyr Glu Glu Leu Leu
           20           25           30
Arg Lys Val Thr Gly Ala Asp Leu Thr Trp Ser Pro Gly Asp Gly
           35           40           45
Ile Gln Phe Gln Val Pro Gly Thr Arg Lys Thr Lys Gln Tyr Cys
           50           55           60
Glu Phe Glu Asn Glu Ile Asn Phe Ile Met Pro His Met Lys Ile
           65           70           75
Gln Ser Leu Leu Phe Leu Leu Gly Phe Tyr Val Lys Asp Pro Ser
           80           85           90
Gln

```

<210> 8  
 <211> 160  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1694122CD1

```

<400> 8
Met Pro Lys Arg Trp Arg Cys Ile Leu Ala Pro Ser Arg Pro Trp
 1           5           10           15
Arg Ser Met Thr Trp Arg Gly Ile Tyr Trp Ile Leu Glu Pro Arg
           20           25           30
Cys Lys Glu Phe Met Gly Ile Met Thr Leu Gly Cys Leu Pro Thr
           35           40           45
Pro Ala Pro Leu His Leu Phe Phe Ser Leu Ser Pro Ala Arg Val
           50           55           60
Leu Arg Ala Pro Tyr Gly Ala Gln Glu Lys Lys Gly Arg Arg Val
           65           70           75
Arg Thr Thr Pro Trp Arg Arg Pro Pro Trp Arg Thr Ser Gly His
           80           85           90
Trp Gly Arg Asp Pro Ile Arg Glu Asn Cys Pro Gln Gln Ser Glu
           95          100          105
Glu Leu Ser Trp Pro Trp Ile Leu Arg Trp Ala Leu Leu Cys Ala
          110          115          120
Leu Arg Gln Ala Thr Cys Pro Leu Ser Leu Ser Phe Leu Ile Cys

```



<223> Incyte ID No: 2886225CD1

<400> 11

Met	Asp	Arg	Trp	Gly	Gln	Asn	Gly	Leu	Phe	Pro	Arg	Arg	Arg	His
1				5					10					15
Leu	Phe	Ala	Pro	Phe	Leu	Asn	Leu	Ile	Ser	Ser	Val	Phe	Leu	His
				20					25					30
Arg	Phe	Cys	Thr	Leu	Gly	Thr	Lys	Lys	Pro	Ser	Gly	Thr	Leu	Leu
				35					40					45
Arg	Lys	Asp	Cys	Arg	Arg	Glu	Asp	Gln	Arg	Glu	Ile	Tyr	Lys	Tyr
				50					55					60
Phe	Arg	Asp	His	Gly	Ile	Tyr	Ser	Arg	Gly	Asn				
				65					70					

<210> 12

<211> 100

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6144418CD1

<400> 12

Met	Asn	Ser	Ala	Val	Gly	Gly	Leu	Ser	Arg	Pro	Phe	Ser	Val	Pro
1				5					10					15
Leu	Thr	Phe	Ser	Ala	Leu	Ile	Pro	Ser	Leu	Leu	Leu	His	Ala	Ser
				20					25					30
Val	Leu	Phe	Cys	Thr	Gly	Trp	Tyr	His	Asp	Phe	Gln	Glu	Gly	Glu
				35					40					45
Ser	Lys	Arg	Glu	Thr	Ser	Gln	Leu	Lys	Gln	Lys	His	Pro	Gly	Thr
				50					55					60
Arg	Glu	Asp	Glu	Val	Asn	Asn	Asp	Ser	Met	Trp	Asp	Thr	Ile	Ser
				65					70					75
His	Cys	His	Ser	Ala	Cys	Ser	Ser	Thr	Asn	Lys	Thr	Ile	Leu	Thr
				80					85					90
Lys	His	Pro	Trp	Ile	Ile	Gly	Ser	His	Asp					
				95					100					

<210> 13

<211> 122

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6834184CD1

<400> 13

Met	Gln	Gly	Val	Pro	Cys	Leu	Gly	Trp	Leu	Leu	Ser	Ser	Ala	Phe
1				5					10					15
Ser	Leu	Met	Ser	Trp	Gly	Ser	Leu	His	Gly	Cys	Ala	Leu	Leu	Leu
				20					25					30
Ala	Leu	Cys	Ser	Gly	Thr	Phe	Glu	Val	Glu	Lys	Ile	Leu	Val	Gly
				35					40					45
Val	Gly	Ala	Asp	Glu	Cys	Gln	Ala	Ser	Ala	Leu	Val	Trp	Glu	Ala
				50					55					60
Thr	Met	Leu	Thr	Phe	Gln	Leu	His	Pro	Arg	Gly	Ser	Thr	Ser	Gln
				65					70					75
Pro	Pro	Glu	Pro	Asp	Cys	Ser	Ala	Ala	Val	Leu	Gly	Lys	Leu	Leu
				80					85					90
Thr	Phe	Leu	Cys	Leu	Ser	Phe	Phe	Ile	Cys	Glu	Leu	Gly	Val	Ile
				95					100					105

Ala Ser Asn Glu Ser Lys Gly Leu Gly Thr Val Thr Lys Leu Trp  
 110 115 120  
 Leu Val

<210> 14  
 <211> 113  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6951005CD1

<400> 14  
 Met Val Leu Pro Gly Phe Pro Ser Val Pro Ser Pro Leu Pro His  
 1 5 10 15  
 Pro Leu Trp Leu Leu Pro Leu Ala Pro Ser Ile Leu Asp Gln Phe  
 20 25 30  
 Ser Leu Gly Pro Thr Leu Arg Ser Pro Ala Phe Ile Pro Ser Arg  
 35 40 45  
 Asp Ser Pro Ala Ser Ile Ala Val Thr Asp Ile Thr Ile His Ile  
 50 55 60  
 Gln Ile Val Leu Leu Ala Thr Leu Leu Ala Ser Ser Phe Thr Lys  
 65 70 75  
 Ser Pro Asp Phe Ser Tyr Asn Pro Asp Leu Ser Phe Thr Ser Ser  
 80 85 90  
 Tyr Met Thr Ser Gly Met Leu Leu Asp Ile Ser Glu Leu Gln Tyr  
 95 100 105  
 Pro Tyr Val Gln Ser Glu Thr Ile  
 110

<210> 15  
 <211> 85  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7250331CD1

<400> 15  
 Met Trp Pro Glu Pro Pro Leu Gly Pro Leu Ser Pro Leu Leu Cys  
 1 5 10 15  
 Leu Leu Ser Leu Ser Cys Leu Pro Glu Val Arg Leu Phe Arg Gly  
 20 25 30  
 Gln Cys Val Thr Cys Gln Leu Pro His His Pro Pro Pro Ser Leu  
 35 40 45  
 Pro Pro Leu Leu Pro Gln Gly Pro Pro Pro Ile Ser Gly Ser Gln  
 50 55 60  
 Ala Ile Asn Leu Glu Thr Glu Met Gly Leu Leu Ser Ile Leu Trp  
 65 70 75  
 Pro Leu Phe Leu Ser Leu Gln Phe Val Pro  
 80 85

<210> 16  
 <211> 256  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1758413CD1

&lt;400&gt; 16

```

Met Ala Pro Gly Ser Arg Thr Ser Leu Leu Leu Ala Phe Ala Leu
 1          5          10          15
Leu Cys Leu Pro Trp Leu Gln Glu Ala Gly Ala Val Gln Thr Val
          20          25          30
Pro Leu Ser Arg Leu Phe Asp His Ala Met Leu Gln Ala His Arg
          35          40          45
Ala His Gln Leu Ala Ile Asp Thr Tyr Gln Glu Phe Glu Glu Thr
          50          55          60
Tyr Ile Pro Lys Asp Gln Lys Tyr Ser Phe Leu His Asp Ser Gln
          65          70          75
Thr Ser Phe Cys Phe Ser Asp Ser Ile Pro Thr Pro Ser Asn Met
          80          85          90
Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser
          95          100          105
Leu Leu Leu Ile Glu Ser Trp Leu Glu Pro Val Arg Phe Leu Arg
          110          115          120
Ser Met Phe Ala Asn Asn Leu Val Tyr Asp Thr Ser Asp Ser Asp
          125          130          135
Asp Tyr His Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu
          140          145          150
Met Gly Val Arg Val Ala Pro Gly Val Thr Asn Pro Gly Thr Pro
          155          160          165
Leu Ala Ser Arg Ala Gly Gly Glu Lys Tyr Cys Cys Pro Leu Phe
          170          175          180
Ser Ser Lys Ala Leu Thr Gln Glu Asn Ser Pro Tyr Ser Ser Phe
          185          190          195
Arg Leu Val Asn Pro Pro Gly Leu Ser Leu His Pro Glu Gly Glu
          200          205          210
Gly Gly Lys Trp Ile Asn Glu Arg Gly Arg Glu Gln Cys Pro Ser
          215          220          225
Ala Trp Pro Leu Leu Leu Phe Leu His Phe Ala Glu Ala Gly Arg
          230          235          240
Arg Gln Pro Pro Asp Trp Ala Asp Pro Gln Ala Asp Leu Gln Gln
          245          250          255
Val

```

&lt;210&gt; 17

&lt;211&gt; 287

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7011042CD1

&lt;400&gt; 17

```

Met Arg Gln Thr Leu Pro Leu Leu Leu Leu Thr Val Leu Arg Pro
 1          5          10          15
Ser Trp Ala Asp Pro Pro Gln Glu Lys Val Pro Leu Phe Arg Val
          20          25          30
Thr Gln Gln Gly Pro Trp Gly Ser Ser Gly Ser Asn Ala Thr Asp
          35          40          45
Ser Pro Cys Glu Gly Leu Pro Ala Ala Asp Ala Thr Ala Leu Thr
          50          55          60
Leu Ala Asn Arg Asn Leu Glu Arg Leu Pro Gly Cys Leu Pro Arg
          65          70          75
Thr Leu Arg Ser Leu Asp Ala Ser His Asn Leu Leu Arg Ala Leu
          80          85          90
Ser Thr Ser Glu Leu Gly His Leu Glu Gln Leu Gln Val Leu Thr
          95          100          105
Leu Arg His Asn Arg Ile Ala Ala Leu Arg Trp Gly Pro Gly Gly

```



110	115	120
Pro Ala Gly Leu His Thr Leu Asp Leu Ser Tyr Asn Gln Leu Ala		
125	130	135
Ala Leu Pro Pro Cys Thr Gly Pro Ala Leu Ser Ser Leu Arg Ala		
140	145	150
Leu Ala Leu Ala Gly Asn Pro Leu Arg Ala Leu Gln Pro Arg Ala		
155	160	165
Phe Ala Cys Phe Pro Ala Leu Gln Leu Leu Asn Leu Ser Cys Thr		
170	175	180
Ala Leu Gly Arg Gly Ala Gln Gly Gly Ile Ala Glu Ala Ala Phe		
185	190	195
Ala Gly Glu Asp Gly Ala Pro Leu Val Thr Leu Glu Val Leu Asp		
200	205	210
Leu Ser Gly Thr Phe Leu Glu Arg Val Glu Ser Gly Trp Ile Arg		
215	220	225
Asp Leu Pro Lys Leu Thr Ser Leu Tyr Leu Arg Lys Met Pro Arg		
230	235	240
Leu Thr Thr Leu Glu Gly Asp Ile Phe Lys Met Thr Pro Asn Leu		
245	250	255
Gln Gln Leu Asp Cys Gln Asp Ser Pro Ala Leu Ala Ser Val Ala		
260	265	270
Thr His Ile Phe Gln Asp Thr Pro His Leu Gln Val Leu Leu Phe		
275	280	285
Gln Lys		

&lt;210&gt; 18

&lt;211&gt; 366

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7427362CD1

&lt;400&gt; 18

Met Leu Asp Gly Ser Pro Leu Ala Arg Trp Leu Ala Ala Ala Phe	
1 5 10 15	
Gly Leu Thr Leu Leu Leu Ala Ala Leu Arg Pro Ser Ala Ala Tyr	
20 25 30	
Phe Gly Leu Thr Gly Ser Glu Pro Leu Thr Ile Leu Pro Leu Thr	
35 40 45	
Leu Glu Pro Glu Ala Ala Ala Gln Ala His Tyr Lys Ala Cys Asp	
50 55 60	
Arg Leu Lys Leu Glu Arg Lys Gln Arg Arg Met Cys Arg Arg Asp	
65 70 75	
Pro Gly Val Ala Glu Thr Leu Val Glu Ala Val Ser Met Ser Ala	
80 85 90	
Leu Glu Cys Gln Phe Gln Phe Arg Phe Glu Arg Trp Asn Cys Thr	
95 100 105	
Leu Glu Gly Arg Tyr Arg Ala Ser Leu Leu Lys Arg Gly Phe Lys	
110 115 120	
Glu Thr Ala Phe Leu Tyr Ala Ile Ser Ser Ala Gly Leu Thr His	
125 130 135	
Ala Leu Ala Lys Ala Cys Ser Ala Gly Arg Met Glu Arg Cys Thr	
140 145 150	
Cys Asp Glu Ala Pro Asp Leu Glu Asn Arg Glu Ala Trp Gln Trp	
155 160 165	
Gly Gly Cys Ser Glu Asp Ile Glu Phe Gly Gly Met Val Ser Arg	
170 175 180	
Glu Phe Ala Asp Ala Arg Glu Asn Arg Pro Asp Ala Arg Ser Ala	
185 190 195	
Met Asn Arg His Asn Asn Glu Ala Gly Arg Gln Val Ile Lys Ala	

	200		205		210
Gly Val Glu Thr	Thr Cys Lys Cys His	Gly Val Ser Gly Ser	Cys		
	215		220		225
Thr Val Arg Thr	Cys Trp Arg Gln Leu	Ala Pro Phe His Glu Val			
	230		235		240
Gly Lys His Leu	Lys His Lys Tyr Glu	Thr Ala Leu Lys Val	Gly		
	245		250		255
Ser Thr Thr Asn	Glu Ala Ala Gly Glu	Ala Gly Ala Ile Ser	Pro		
	260		265		270
Pro Arg Gly Arg	Ala Ser Gly Ala Gly	Gly Ser Asp Pro Leu	Pro		
	275		280		285
Arg Thr Pro Glu	Leu Val His Leu Asp	Asp Ser Pro Ser Phe	Cys		
	290		295		300
Leu Ala Gly Arg	Phe Ser Pro Gly Thr	Ala Gly Arg Arg Cys	His		
	305		310		315
Arg Glu Lys Asn	Cys Glu Ser Ile Cys	Cys Gly Arg Gly His	Asn		
	320		325		330
Thr Gln Ser Arg	Val Val Thr Arg Pro	Cys Gln Cys Gln Val	Arg		
	335		340		345
Trp Cys Cys Tyr	Val Glu Cys Arg Gln	Cys Thr Gln Arg Glu	Glu		
	350		355		360
Val Tyr Thr Cys	Lys Gly				
	365				

<210> 19  
 <211> 416  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7485304CD1

<400> 19

Met Leu Ala Val Val	Met Ala Asp Leu Ala	Ser Leu Met Cys Trp	
1	5	10	15
Val Cys Lys Gln Lys	Leu Pro Gly Leu Ala	Ala Trp Ser Ala Ala	
	20	25	30
Val Arg Gln Glu Val	Gly Leu Cys Leu Glu	Arg Gln Ser Leu Gln	
	35	40	45
Leu Asp Pro Ala Leu	Ser Ser Leu Ser Gln	Gly Trp Pro Leu Arg	
	50	55	60
Arg Pro Leu Pro Phe	Ile Cys Pro Ser Pro	Pro Ser Pro Arg Leu	
	65	70	75
Thr Cys Leu Pro Pro	Leu Ala Leu Ser Ser	Leu Thr Gly Arg Glu	
	80	85	90
Val Leu Thr Pro Phe	Pro Gly Leu Gly Thr	Ala Ala Ala Pro Ala	
	95	100	105
Gln Gly Gly Ala His	Leu Lys Gln Cys Asp	Leu Leu Lys Leu Ser	
	110	115	120
Arg Arg Gln Lys Gln	Leu Cys Arg Arg Glu	Pro Gly Leu Ala Glu	
	125	130	135
Thr Leu Arg Asp Ala	Ala His Leu Gly Leu	Leu Glu Cys Gln Phe	
	140	145	150
Gln Phe Arg His Glu	Arg Trp Asn Cys Ser	Leu Glu Gly Arg Met	
	155	160	165
Gly Leu Leu Lys Arg	Gly Phe Lys Glu Thr	Ala Phe Leu Tyr Ala	
	170	175	180
Val Ser Ser Ala Ala	Leu Thr His Thr Leu	Ala Arg Ala Cys Ser	
	185	190	195
Ala Gly Arg Met Glu	Arg Cys Thr Cys Asp	Asp Ser Pro Gly Leu	
	200	205	210
Glu Ser Arg Gln Ala	Trp Gln Trp Gly Val	Cys Gly Asp Asn Leu	

	215		220		225
Lys Tyr Ser Thr	Lys Phe Leu Ser Asn	Phe Leu Gly Ser Lys	Arg		
	230		235		240
Gly Asn Lys Asp	Leu Arg Ala Arg Ala	Asp Ala His Asn Thr	His		
	245		250		255
Val Gly Ile Lys	Ala Val Lys Ser Gly	Leu Arg Thr Thr Cys	Lys		
	260		265		270
Cys His Gly Val	Ser Gly Ser Cys Ala	Val Arg Thr Cys Trp	Lys		
	275		280		285
Gln Leu Ser Pro	Phe Arg Glu Thr Gly	Gln Val Leu Lys Leu	Arg		
	290		295		300
Tyr Asp Ser Ala	Val Lys Val Ser Ser	Ala Thr Asn Glu Ala	Leu		
	305		310		315
Gly Arg Leu Glu	Leu Trp Ala Pro Ala	Arg Gln Gly Ser Leu	Thr		
	320		325		330
Lys Gly Leu Ala	Pro Arg Ser Gly Asp	Leu Val Tyr Met Glu	Asp		
	335		340		345
Ser Pro Ser Phe	Cys Arg Pro Ser Lys	Tyr Ser Pro Gly Thr	Ala		
	350		355		360
Gly Arg Val Cys	Ser Arg Glu Ala Ser	Cys Ser Ser Leu Cys	Cys		
	365		370		375
Gly Arg Gly Tyr	Asp Thr Gln Ser Arg	Leu Val Ala Phe Ser	Cys		
	380		385		390
His Cys Gln Val	Gln Trp Cys Cys Tyr	Val Glu Cys Gln Gln	Cys		
	395		400		405
Val Gln Glu Glu	Leu Val Tyr Thr Cys	Lys His			
	410		415		

&lt;210&gt; 20

&lt;211&gt; 871

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1422394CD1

&lt;400&gt; 20

Met Lys Tyr Ser Cys	Cys Ala Leu Val Leu	Ala Val Leu Gly Thr	
1	5	10	15
Glu Leu Leu Gly Ser	Leu Cys Ser Thr Val	Arg Ser Pro Arg Phe	
	20	25	30
Arg Gly Arg Ile Gln	Gln Glu Arg Lys Asn	Ile Arg Pro Asn Ile	
	35	40	45
Ile Leu Val Leu Thr	Asp Asp Gln Asp Val	Glu Leu Gly Ser Leu	
	50	55	60
Gln Val Met Asn Lys	Thr Arg Lys Ile Met	Glu His Gly Gly Ala	
	65	70	75
Thr Phe Ile Asn Ala	Phe Val Thr Thr Pro	Met Cys Cys Pro Ser	
	80	85	90
Arg Ser Ser Met Leu	Thr Gly Lys Tyr Val	His Asn His Asn Val	
	95	100	105
Tyr Thr Asn Asn Glu	Asn Cys Ser Ser Pro	Ser Trp Gln Ala Met	
	110	115	120
His Glu Pro Arg Thr	Phe Ala Val Tyr Leu	Asn Asn Thr Gly Tyr	
	125	130	135
Arg Thr Ala Phe Phe	Gly Lys Tyr Leu Asn	Glu Tyr Asn Gly Ser	
	140	145	150
Tyr Ile Pro Pro Gly	Trp Arg Glu Trp Leu	Gly Leu Ile Lys Asn	
	155	160	165
Ser Arg Phe Tyr Asn	Tyr Thr Val Cys Arg	Asn Gly Ile Lys Glu	
	170	175	180
Lys His Gly Phe Asp	Tyr Ala Lys Asp Tyr	Phe Thr Asp Leu Ile	

	185		190		195
Thr Asn Glu Ser	Ile Asn Tyr Phe Lys	Met Ser Lys Arg Met	Tyr		
	200		205		210
Pro His Arg Pro	Val Met Met Val Ile	Ser His Ala Ala Pro	His		
	215		220		225
Gly Pro Glu Asp	Ser Ala Pro Gln Phe	Ser Lys Leu Tyr Pro	Asn		
	230		235		240
Ala Ser Gln His	Ile Thr Pro Ser Tyr	Asn Tyr Ala Pro Asn	Met		
	245		250		255
Asp Lys His Trp	Ile Met Gln Tyr Thr	Gly Pro Met Leu Pro	Ile		
	260		265		270
His Met Glu Phe	Thr Asn Ile Leu Gln	Arg Lys Arg Leu Gln	Thr		
	275		280		285
Leu Met Ser Val	Asp Asp Ser Val Glu	Arg Leu Tyr Asn Met	Leu		
	290		295		300
Val Glu Thr Gly	Glu Leu Glu Asn Thr	Tyr Ile Ile Tyr Thr	Ala		
	305		310		315
Asp His Gly Tyr	His Ile Gly Gln Phe	Gly Leu Val Lys Gly	Lys		
	320		325		330
Ser Met Pro Tyr	Asp Phe Asp Ile Arg	Val Pro Phe Phe Ile	Arg		
	335		340		345
Gly Pro Ser Val	Glu Pro Gly Ser Ile	Val Pro Gln Ile Val	Leu		
	350		355		360
Asn Ile Asp Leu	Ala Pro Thr Ile Leu	Asp Ile Ala Gly Leu	Asp		
	365		370		375
Thr Pro Pro Asp	Val Asp Gly Lys Ser	Val Leu Lys Leu Leu	Asp		
	380		385		390
Pro Glu Lys Pro	Gly Asn Arg Phe Arg	Thr Asn Lys Lys Ala	Lys		
	395		400		405
Ile Trp Arg Asp	Thr Phe Leu Val Glu	Arg Gly Lys Phe Leu	Arg		
	410		415		420
Lys Lys Glu Glu	Ser Ser Lys Asn Ile	Gln Gln Ser Asn His	Leu		
	425		430		435
Pro Lys Tyr Glu	Arg Val Lys Glu Leu	Cys Gln Gln Ala Arg	Tyr		
	440		445		450
Gln Thr Ala Cys	Glu Gln Pro Gly Gln	Lys Trp Gln Cys Ile	Glu		
	455		460		465
Asp Thr Ser Gly	Lys Leu Arg Ile His	Lys Cys Lys Gly Pro	Ser		
	470		475		480
Asp Leu Leu Thr	Val Arg Gln Ser Thr	Arg Asn Leu Tyr Ala	Arg		
	485		490		495
Gly Phe His Asp	Lys Asp Lys Glu Cys	Ser Cys Arg Glu Ser	Gly		
	500		505		510
Tyr Arg Ala Ser	Arg Ser Gln Arg Lys	Ser Gln Arg Gln Phe	Leu		
	515		520		525
Arg Asn Gln Gly	Thr Pro Lys Tyr Lys	Pro Arg Phe Val His	Thr		
	530		535		540
Arg Gln Thr Arg	Ser Leu Ser Val Glu	Phe Glu Gly Glu Ile	Tyr		
	545		550		555
Asp Ile Asn Leu	Glu Glu Glu Glu Glu	Leu Gln Val Leu Gln	Pro		
	560		565		570
Arg Asn Ile Ala	Lys Arg His Asp Glu	Gly His Lys Gly Pro	Arg		
	575		580		585
Asp Leu Gln Ala	Ser Ser Gly Gly Asn	Arg Gly Arg Met Leu	Ala		
	590		595		600
Asp Ser Ser Asn	Ala Val Gly Pro Pro	Thr Thr Val Arg Val	Thr		
	605		610		615
His Lys Cys Phe	Ile Leu Pro Asn Asp	Ser Ile His Cys Glu	Arg		
	620		625		630
Glu Leu Tyr Gln	Ser Ala Arg Ala Trp	Lys Asp His Lys Ala	Tyr		
	635		640		645
Ile Asp Lys Glu	Ile Glu Ala Leu Gln	Asp Lys Ile Lys Asn	Leu		
	650		655		660

Arg	Glu	Val	Arg	Gly	His	Leu	Lys	Arg	Arg	Lys	Pro	Glu	Glu	Cys	
				665					670					675	
Ser	Cys	Ser	Lys	Gln	Ser	Tyr	Tyr	Asn	Lys	Glu	Lys	Gly	Val	Lys	
				680					685					690	
Lys	Gln	Glu	Lys	Leu	Lys	Ser	His	Leu	His	Pro	Phe	Lys	Glu	Ala	
				695					700					705	
Ala	Gln	Glu	Val	Asp	Ser	Lys	Leu	Gln	Leu	Phe	Lys	Glu	Asn	Asn	
				710					715					720	
Arg	Arg	Arg	Lys	Lys	Glu	Arg	Lys	Glu	Lys	Arg	Arg	Gln	Arg	Lys	
				725					730					735	
Gly	Glu	Glu	Cys	Ser	Leu	Pro	Gly	Leu	Thr	Cys	Phe	Thr	His	Asp	
				740					745					750	
Asn	Asn	His	Trp	Gln	Thr	Ala	Pro	Phe	Trp	Asn	Leu	Gly	Ser	Phe	
				755					760					765	
Cys	Ala	Cys	Thr	Ser	Ser	Asn	Asn	Asn	Thr	Tyr	Trp	Cys	Leu	Arg	
				770					775					780	
Thr	Val	Asn	Glu	Thr	His	Asn	Phe	Leu	Phe	Cys	Glu	Phe	Ala	Thr	
				785					790					795	
Gly	Phe	Leu	Glu	Tyr	Phe	Asp	Met	Asn	Thr	Asp	Pro	Tyr	Gln	Leu	
				800					805					810	
Thr	Asn	Thr	Val	His	Thr	Val	Glu	Arg	Gly	Ile	Leu	Asn	Gln	Leu	
				815					820					825	
His	Val	Gln	Leu	Met	Glu	Leu	Arg	Ser	Cys	Gln	Gly	Tyr	Lys	Gln	
				830					835					840	
Cys	Asn	Pro	Arg	Pro	Lys	Asn	Leu	Asp	Val	Gly	Asn	Lys	Asp	Gly	
				845					850					855	
Gly	Ser	Tyr	Asp	Leu	His	Arg	Gly	Gln	Leu	Trp	Asp	Gly	Trp	Glu	
				860					865					870	
Gly															

<210> 21  
 <211> 100  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1336022CD1

<400> 21																
Met	Lys	Ser	Val	Asn	Asp	Thr	Leu	Leu	Ala	His	Phe	Leu	Val	Leu		
1				5					10					15		
Leu	Val	Ile	Leu	Pro	Pro	Ala	Pro	Val	Lys	Pro	Val	Pro	Gly	His		
				20					25					30		
Ile	Thr	Gln	Leu	Pro	Ala	Gln	Leu	Leu	Arg	Glu	Lys	Thr	Met	His		
				35					40					45		
Phe	Thr	Ser	Thr	Ser	Pro	Ala	Thr	Gly	Thr	Gln	Met	Val	Asn	Ala		
				50					55					60		
Ala	Ala	Asn	Gly	Leu	Gly	Ala	Glu	Pro	Met	Glu	Ser	Phe	Lys	Gln		
				65					70					75		
Ala	Tyr	Arg	His	Cys	Ile	Lys	Ile	Pro	Asp	Phe	Lys	Ile	Pro	Ser		
				80					85					90		
Gln	Gly	Ser	His	Lys	Thr	Ile	Ile	Phe	Ser							
				95					100							

<210> 22  
 <211> 102  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature

<400> 22

<210> 23

<211> 117

<212> PRT

<213> Homo sapiens

<220>

```
<221> misc_feature
```

<223> Incyte ID No: 7475846CD1

<400> 23

Met	Cys	His	Gly	Ser	Pro	Thr	Leu	Cys	Gln	Pro	Val	Cys	Ala	Met
1				5					10					15
Ala	Pro	Asp	Pro	Val	Pro	Ala	His	Val	Cys	His	Gly	Ser	Pro	Thr
				20					25					30
Leu	Cys	Gln	Pro	Val	Trp	Ala	Met	Ala	Pro	Pro	Asn	Pro	Cys	Gln
				35					40					45
Pro	Ala	Cys	Ala	Met	Gly	Ser	Thr	Asp	Pro	Val	Pro	Ala	Arg	Val
				50					55					60
Arg	His	Gly	Phe	Pro	Asp	Pro	Met	Pro	Ala	Arg	Val	Cys	Ala	Met
				65					70					75
Ala	Pro	Pro	Thr	Pro	Cys	Gln	Pro	Ala	Cys	Val	Met	Thr	Pro	Pro
				80					85					90
Arg	Val	Arg	His	Gly	Phe	Pro	Asp	Pro	Met	Pro	Ala	Arg	Val	Arg
				95					100					105
His	Gly	Ser	Thr	Asp	Pro	Val	Pro	Ala	Ser	Ala	Gly			
				110					115					

<210> 24

<211> 150

<212> PRT

<213> Homo sapiens

<220>

```
<221> misc_feature
```

<223> Incyte ID No: 7475860CD1

<400> 24

Met	Ala	Ala	Ser	Gln	Cys	Leu	Cys	Cys	Ser	Lys	Phe	Leu	Phe	Gln
1				5					10					15
Arg	Gln	Asn	Leu	Ala	Cys	Phe	Leu	Thr	Asn	Pro	His	Cys	Gly	Ser
				20					25					30
Leu	Val	Asn	Ala	Asp	Gly	His	Gly	Glu	Val	Trp	Thr	Asp	Trp	Asn
				35					40					45
Asn	Met	Ser	Lys	Phe	Phe	Gln	Tyr	Gly	Trp	Arg	Cys	Thr	Thr	Asn
				50					55					60

Glu	Asn	Thr	Tyr	Ser	Asn	Arg	Thr	Leu	Met	Gly	Asn	Trp	Asn	Gln
				65					70					75
Glu	Arg	Tyr	Asp	Leu	Arg	Asn	Ile	Val	Gln	Pro	Lys	Pro	Leu	Pro
				80					85					90
Ser	Gln	Phe	Gly	His	Tyr	Phe	Glu	Thr	Thr	Tyr	Asp	Thr	Ser	Tyr
				95					100					105
Asn	Asn	Lys	Met	Pro	Leu	Ser	Thr	His	Arg	Phe	Lys	Arg	Glu	Pro
				110					115					120
His	Trp	Phe	Pro	Gly	His	Gln	Pro	Glu	Leu	Asp	Pro	Pro	Arg	Tyr
				125					130					135
Lys	Cys	Thr	Glu	Lys	Ser	Thr	Tyr	Met	Asn	Ser	Tyr	Ser	Lys	Pro
				140					145					150

&lt;210&gt; 25

&lt;211&gt; 89

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7950941CD1

&lt;400&gt; 25

Met	Ala	Pro	Asn	Pro	Ala	Arg	Leu	His	Ser	His	Leu	Asp	Leu	Val
1				5					10					15
Ser	Pro	Ser	Val	Pro	Arg	Ser	Leu	Gly	Phe	Gln	Leu	Pro	Ile	Gly
				20					25					30
Arg	Lys	Gln	Ser	Arg	Asn	Val	Leu	Ser	His	Gln	Asp	Gly	His	Ile
				35					40					45
Leu	Gln	Cys	Ser	Phe	Arg	Pro	Asp	Arg	Arg	Met	Lys	Arg	Lys	Ala
				50					55					60
Glu	Ser	Pro	Glu	Asn	Asn	Gln	Leu	Arg	Cys	His	Leu	Pro	Cys	Gln
				65					70					75
Gly	Gly	Asp	Pro	Ala	Met	Leu	Pro	Ser	Arg	Phe	Gln	Asn	Cys	
				80					85					

&lt;210&gt; 26

&lt;211&gt; 287

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7485334CD1

&lt;400&gt; 26

Met	Ala	Leu	Gly	Leu	Leu	Ile	Ala	Val	Pro	Leu	Leu	Leu	Gln	Ala
1				5					10					15
Ala	Pro	Pro	Gly	Ala	Ala	His	Tyr	Glu	Met	Leu	Gly	Thr	Cys	Arg
				20					25					30
Met	Ile	Cys	Asp	Pro	Tyr	Ser	Val	Ala	Pro	Ala	Gly	Gly	Pro	Ala
				35					40					45
Gly	Ala	Lys	Ala	Pro	Pro	Pro	Gly	Pro	Ser	Thr	Ala	Ala	Leu	Glu
				50					55					60
Val	Met	Gln	Asp	Leu	Ser	Ala	Asn	Pro	Pro	Pro	Pro	Phe	Ile	Gln
				65					70					75
Gly	Pro	Lys	Gly	Asp	Pro	Gly	Arg	Pro	Gly	Lys	Pro	Gly	Pro	Arg
				80					85					90
Gly	Pro	Pro	Gly	Glu	Pro	Gly	Pro	Pro	Gly	Pro	Arg	Gly	Pro	Pro
				95					100					105
Gly	Glu	Lys	Gly	Asp	Ser	Gly	Arg	Pro	Gly	Leu	Pro	Gly	Leu	Gln
				110					115					120

Leu Thr Thr Ser	Ala Ala Gly Gly Val	Gly Val Val Ser Gly Gly	
	125	130	135
Thr Gly Gly Gly	Gly Asp Thr Glu Gly	Glu Val Thr Ser Ala Leu	
	140	145	150
Ser Ala Ala Phe	Ser Gly Pro Lys Ile	Ala Phe Tyr Val Gly Leu	
	155	160	165
Lys Ser Pro His	Glu Gly Tyr Glu Val	Leu Lys Phe Asp Asp Val	
	170	175	180
Val Thr Asn Leu	Gly Asn His Tyr Asp	Pro Thr Thr Gly Lys Phe	
	185	190	195
Ser Cys Gln Val	Arg Gly Ile Tyr Phe	Phe Thr Tyr His Ile Leu	
	200	205	210
Met Arg Gly Gly	Asp Gly Thr Ser Met	Trp Ala Asp Leu Cys Lys	
	215	220	225
Asn Gly Gln Val	Arg Ala Ser Ala Ile	Ala Gln Asp Ala Asp Gln	
	230	235	240
Asn Tyr Asp Tyr	Ala Ser Asn Ser Val	Val Leu His Leu Asp Ser	
	245	250	255
Gly Asp Glu Val	Tyr Val Lys Leu Asp	Gly Gly Lys Ala His Gly	
	260	265	270
Gly Asn Asn Asn	Lys Tyr Ser Thr Phe	Ser Gly Phe Leu Leu Tyr	
	275	280	285
Pro Asp			

<210> 27  
 <211> 578  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7220001CD1

<400> 27

Met Asp Gly Glu Ala Thr Val Lys Pro Gly Glu Gln Lys Glu Val	
1 5 10 15	
Val Arg Arg Gly Arg Glu Val Asp Tyr Ser Arg Leu Ile Ala Gly	
20 25 30	
Thr Leu Pro Gln Ser His Val Thr Ser Arg Arg Ala Gly Trp Lys	
35 40 45	
Met Pro Leu Phe Leu Ile Leu Cys Leu Leu Gln Gly Ser Ser Phe	
50 55 60	
Ala Leu Pro Gln Lys Arg Pro His Pro Arg Trp Leu Trp Glu Gly	
65 70 75	
Ser Leu Pro Ser Arg Thr His Leu Arg Ala Met Gly Thr Leu Arg	
80 85 90	
Pro Ser Ser Pro Leu Cys Trp Arg Glu Glu Ser Ser Phe Ala Ala	
95 100 105	
Pro Asn Ser Leu Lys Gly Ser Arg Leu Val Ser Gly Glu Pro Gly	
110 115 120	
Gly Ala Val Thr Ile Gln Cys His Tyr Ala Pro Ser Ser Val Asn	
125 130 135	
Arg His Gln Arg Lys Tyr Trp Cys Arg Leu Gly Pro Pro Arg Trp	
140 145 150	
Ile Cys Gln Thr Ile Val Ser Thr Asn Gln Tyr Thr His His Arg	
155 160 165	
Tyr Arg Asp Arg Val Ala Leu Thr Asp Phe Pro Gln Arg Gly Leu	
170 175 180	
Phe Val Val Arg Leu Ser Gln Leu Ser Pro Asp Asp Ile Gly Cys	
185 190 195	
Tyr Leu Cys Gly Ile Gly Ser Glu Asn Asn Met Leu Phe Leu Ser	
200 205 210	



Met Asn Leu Thr	Ile Ser Ala Gly Pro	Ala Ser Thr Leu Pro	Thr
215	220		225
Ala Thr Pro Ala	Ala Gly Glu Leu Thr	Met Arg Ser Tyr Gly	Thr
230	235		240
Ala Ser Pro Val	Ala Asn Arg Trp Thr	Pro Gly Ser His Pro	Asp
245	250		255
Leu Arg Thr Gly	Asp Ser Met Gly His	Met Leu Leu Pro His	Pro
260	265		270
Gly Thr Ser Lys	Thr Thr Ala Ser Ala	Glu Gly Arg Arg Thr	Pro
275	280		285
Gly Ala Thr Arg	Pro Ala Ala Pro Gly	Thr Gly Ser Trp Ala	Glu
290	295		300
Gly Ser Val Lys	Ala Pro Ala Pro Ile	Pro Glu Ser Pro Pro	Ser
305	310		315
Lys Ser Arg Ser	Met Ser Asn Thr Thr	Glu Gly Val Arg Glu	Gly
320	325		330
Thr Arg Ser Ser	Val Thr Asn Arg Ala	Arg Ala Ser Lys Asp	Arg
335	340		345
Arg Glu Met Thr	Thr Thr Lys Ala Asp	Arg Pro Arg Glu Asp	Ile
350	355		360
Glu Gly Val Arg	Ile Ala Leu Asp Ala	Ala Lys Lys Val Leu	Gly
365	370		375
Thr Ile Gly Pro	Pro Ala Leu Val Ser	Glu Thr Leu Ala Trp	Glu
380	385		390
Ile Leu Pro Gln	Ala Thr Pro Val Ser	Lys Gln Gln Ser Gln	Gly
395	400		405
Ser Ile Gly Glu	Thr Thr Pro Ala Ala	Gly Met Trp Thr Leu	Gly
410	415		420
Thr Pro Ala Ala	Asp Val Trp Ile Leu	Gly Thr Pro Ala Ala	Asp
425	430		435
Val Trp Thr Ser	Met Glu Ala Ala Ser	Gly Glu Gly Ser Ala	Ala
440	445		450
Gly Asp Leu Asp	Ala Ala Thr Gly Asp	Arg Gly Pro Gln Ala	Thr
455	460		465
Leu Ser Gln Thr	Pro Ala Val Gly Pro	Trp Gly Pro Pro Gly	Lys
470	475		480
Glu Ser Ser Val	Lys Arg Thr Phe Pro	Glu Asp Glu Ser Ser	Ser
485	490		495
Arg Thr Leu Ala	Pro Val Ser Thr Met	Leu Ala Leu Phe Met	Leu
500	505		510
Met Ala Leu Val	Leu Leu Gln Arg Lys	Leu Trp Arg Arg Arg	Thr
515	520		525
Ser Gln Glu Ala	Glu Arg Val Thr Leu	Ile Gln Met Thr His	Phe
530	535		540
Leu Glu Val Asn	Pro Gln Ala Asp Gln	Leu Pro His Val Glu	Arg
545	550		555
Lys Met Leu Gln	Asp Asp Ser Leu Pro	Ala Gly Ala Ser Leu	Thr
560	565		570
Ala Pro Glu Arg	Asn Pro Gly Pro		
575			

<210> 28  
 <211> 285  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5956275CD1

<400> 28  
 Met Glu Gln Arg Asn Arg Leu Gly Ala Leu Gly Tyr Leu Pro Pro  
 1 5 10 15

```

Leu Leu Leu His Ala Leu Leu Leu Phe Val Ala Asp Ala Ala Phe
20 25 30
Thr Glu Val Pro Lys Asp Val Thr Val Arg Glu Gly Asp Asp Ile
35 40 45
Glu Met Pro Cys Ala Phe Arg Ala Ser Gly Ala Thr Ser Tyr Ser
50 55 60
Leu Glu Ile Gln Trp Tyr Leu Lys Glu Pro Pro Arg Glu Leu
65 70 75
Leu His Glu Leu Ala Leu Ser Val Pro Gly Ala Arg Ser Lys Val
80 85 90
Thr Asn Lys Asp Ala Thr Lys Ile Ser Thr Val Arg Val Gln Gly
95 100 105
Asn Asp Ile Ser His Arg Leu Arg Leu Ser Ala Val Arg Leu Gln
110 115 120
Asp Glu Gly Val Tyr Glu Cys Arg Val Ser Asp Tyr Ser Asp Asp
125 130 135
Asp Thr Gln Glu His Lys Ala Gln Ala Met Leu Arg Val Leu Ser
140 145 150
Arg Phe Ala Pro Pro Asn Met Gln Ala Ala Glu Ala Val Ser His
155 160 165
Ile Gln Ser Ser Gly Pro Arg Arg His Gly Pro Ala Ser Ala Ala
170 175 180
Asn Ala Asn Asn Ala Gly Ala Ala Ser Arg Thr Thr Ser Glu Pro
185 190 195
Gly Arg Gly Asp Lys Ser Pro Pro Pro Gly Ser Pro Pro Ala Ala
200 205 210
Ile Asp Pro Ala Val Pro Glu Ala Ala Ala Ala Ser Ala Ala His
215 220 225
Thr Pro Thr Thr Thr Val Ala Ala Ala Ala Ala Ala Ser Ser Ala
230 235 240
Ser Pro Pro Ser Gly Gln Ala Val Leu Leu Arg Gln Arg His Gly
245 250 255
Ser Gly Lys Gly Arg Ser Tyr Thr Thr Asp Pro Leu Leu Ser Leu
260 265 270
Leu Leu Leu Ala Leu His Lys Phe Leu Arg Leu Leu Leu Gly His
275 280 285

```

<210> 29  
 <211> 72  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 346472CD1

```

<400> 29
Met Val Phe Ile Phe Phe Leu Phe Ser Gly Cys Leu Leu Cys Phe
1 5 10 15
Ser Phe Leu Gln Ser Asn Phe Gln His Ser Asp Lys Pro Phe Glu
20 25 30
Arg Asn Arg Leu Arg Ile Pro Tyr Ser Gln Asn Cys Gly Ile Phe
35 40 45
Lys Pro Gln Arg Lys Pro Arg Asp Pro Arg Arg Leu Phe Cys Gly
50 55 60
Cys Gly Lys Phe Lys Tyr Pro Pro Arg Leu His Ser
65 70

```

<210> 30  
 <211> 72  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 643526CD1

<400> 30

Met	Thr	Thr	Leu	Tyr	Leu	Pro	Ala	Phe	Ala	Ala	Val	Leu	Ser	Leu	1	5	10	15
Ser	Gln	Cys	Ser	Glu	Ser	Val	Gly	Ser	Phe	Pro	Thr	Gln	Val	Leu	20	25	30	35
Ala	Ala	Asp	Leu	Gly	Leu	Ala	Leu	Leu	Asp	Val	Ile	Leu	Gln	Pro	40	45	50	55
Arg	Gly	Lys	Leu	Ser	Leu	Tyr	Val	Pro	Ser	Thr	Ala	Trp	Gly	Gln	60	65	70	75
Thr	Arg	Thr	Leu	Thr	Val	Ala	Met	Ala	Glu	Gly	Leu							

<210> 31  
 <211> 149  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1483418CD1

<400> 31

Met	Arg	Pro	Thr	Gly	Gly	Ser	Gly	Gln	Arg	Gly	Pro	Arg	Tyr	Thr	1	5	10	15
Thr	Ser	Leu	Leu	Phe	His	Cys	Leu	Leu	Pro	Cys	Ser	Asp	His	Ser	20	25	30	35
Ser	Gly	Ala	Val	Ser	Gln	Ala	Trp	Ala	Ser	Phe	Asn	Ile	Phe	Tyr	40	45	50	55
Leu	Ala	Leu	His	Gly	Ala	Ala	Pro	Ala	Met	Val	Pro	Gln	Gly	Phe	60	65	70	75
Phe	Ser	Gln	Val	Ser	Ser	Leu	Glu	Arg	Ser	Pro	Arg	Phe	Pro	Val	80	85	90	95
Lys	Gln	Pro	Cys	Ser	Leu	Cys	Leu	Ser	Gln	Pro	His	His	Pro	Val	100	105	110	115
Ala	Ser	Phe	Thr	Ala	Cys	Leu	Thr	Ile	Cys	Asn	His	Leu	Ser	Val	120	125	130	135
Cys	Arg	Leu	Val	Asp	Leu	Leu	Pro	Pro	His	Cys	Gln	Leu	Leu	Gly	140	145		
Asn	Arg	Asp	Trp	Phe	Val	Tyr	Cys	Ala	Ser	Leu	Val	Pro	Arg	Thr				
Gly	His	Gly	Ile	Leu	Leu	Val	His	Asn	Lys	Tyr	Gly	Gly	Asn					

<210> 32  
 <211> 100  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2683477CD1

<400> 32

Met	Pro	Phe	Ser	Asn	Pro	Met	Ala	Ser	Ser	Ser	Pro	Ser	Gly	Trp	1	5	10	15
Pro	Arg	Ala	Ala	Gly	Lys	Ala	Leu	Met	Val	Trp	Val	Val	Leu	Phe	20	25	30	35
Pro	Trp	Ala	Glu	Leu	Gly	Trp	Arg	Thr	Leu	Ser	Arg	Val	Ala	Ala	40	45		

Ser	Leu	Trp	Gly	Pro	Tyr	Leu	Gly	Thr	Tyr	Thr	Asp	Gln	Ala	Val	
				50					55					60	
Cys	Leu	Cys	Ser	Leu	Ser	Asn	His	Asn	Tyr	Ser	Gln	Lys	Ala	Cys	
				65					70					75	
Gly	Leu	Glu	Ser	Thr	Thr	Val	Lys	Pro	Gly	Arg	Met	Cys	Tyr	Pro	
				80					85					90	
Val	Pro	Glu	Arg	Leu	Leu	Val	Cys	Val	Leu						
				95					100						

<210> 33  
 <211> 78  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5580991CD1

<400> 33															
Met	Asn	Ile	Met	Pro	Tyr	Leu	Leu	Gln	Leu	Ser	Phe	Phe	Leu	Leu	
1				5					10					15	
Leu	Phe	Ser	Leu	Pro	Phe	Ser	Leu	Cys	Pro	Ser	Ser	Leu	Ser	Leu	
				20					25					30	
Leu	Phe	Phe	Leu	Leu	Ala	Val	Gly	Phe	Tyr	Phe	Phe	Phe	Glu	Thr	
				35					40					45	
Ser	Leu	Ala	Leu	Ser	Pro	Arg	Leu	Glu	Cys	Ser	Gly	Ala	Ile	Ser	
				50					55					60	
Ala	His	Cys	Lys	Leu	Cys	Leu	Pro	Gly	Ser	Cys	Tyr	Ser	Trp	Ala	
				65					70					75	
Ser	Ala	Cys													

<210> 34  
 <211> 75  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5605931CD1

<400> 34															
Met	Gly	Ser	Pro	Ala	Leu	Gln	Met	Cys	Val	Leu	Thr	Leu	Cys	Leu	
1				5					10					15	
Asp	Leu	Phe	Leu	Leu	Gly	Leu	Arg	Thr	Phe	Cys	Pro	Gln	Met	Ser	
				20					25					30	
Pro	Leu	Val	Thr	Val	Cys	Leu	Arg	Ala	Leu	Gly	Leu	Ala	Gly	Trp	
				35					40					45	
Glu	Gln	Thr	Gln	Leu	Cys	Gly	Gly	His	Gln	Val	Val	Pro	Phe	Ile	
				50					55					60	
Ser	Ser	Gly	Leu	Ser	Leu	Leu	Glu	Cys	Gly	Arg	Cys	Gln	Lys	Gln	
				65					70					75	

<210> 35  
 <211> 111  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6975241CD1

<400> 35  
 Met Val Ser Ser Val Ser Ile Arg Gln Ser Gln Val Leu Val Leu  
 1 5 10 15  
 Cys Leu Cys Leu Cys Leu Glu Gln Lys Leu Val Pro Gly Val Ile  
 20 25 30  
 Cys Lys Gln Glu Ile Leu Arg Glu Met Gly Met Trp Glu Asp Thr  
 35 40 45  
 Gly Val Ala Arg Ser Ser Cys Thr Glu Val Asn Lys Asn Pro Ala  
 50 55 60  
 Gly Ser Ser Trp Met Gly Ile Gln Gln Thr Arg Ala His Asn Ser  
 65 70 75  
 Gly Arg Ala Thr Tyr Thr Gly Ala Cys Asp Trp Leu Gln Trp Ser  
 80 85 90  
 Pro Leu Arg Ala Arg Asp Pro Ala Ala Ile Lys Gln Glu Lys Leu  
 95 100 105  
 Gln Val Gly Ser Arg Phe  
 110

<210> 36  
 <211> 72  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6988529CD1

<400> 36  
 Met Gln Ser Leu Leu Leu Leu Gly Ala Val Val Thr Val Ile Ala  
 1 5 10 15  
 Glu Thr Glu Ile Ala Lys Pro Val Leu Tyr Lys Glu Cys Ala Ser  
 20 25 30  
 Ala Ile Glu Asp Thr Ala Arg Ile Gly Cys Trp Ser Ser Ala Gly  
 35 40 45  
 Pro Ala Val Ile Thr Arg Val Gln Gln Arg Glu Ser Pro Pro Leu  
 50 55 60  
 Pro Ser Leu Thr Gln His Leu Thr Leu Ser His Ser  
 65 70

<210> 37  
 <211> 90  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6996808CD1

<400> 37  
 Met Phe Cys Ala Phe Leu Phe Leu Pro Phe Ser Gln Asp Val Leu  
 1 5 10 15  
 Cys Met Cys Phe Gly Lys Val Val Leu Val Met Phe Ile Leu Leu  
 20 25 30  
 Cys Ile Cys Ser Val Leu Glu Leu Phe Phe Ser Ser Gly Arg Cys  
 35 40 45  
 Phe Glu Ser Thr Leu Phe Ile Val Ala His Val Ser Asn Leu Ile  
 50 55 60  
 Ser Lys Ile Leu Gln Val Tyr Ser Leu Arg Arg Ile Leu Phe Ile  
 65 70 75  
 Tyr Cys Thr Asp Met Leu Cys Thr Arg His Cys Ala Met Ala Asn  
 80 85 90

<210> 38  
 <211> 283  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7472689CD1

<400> 38  
 Met Trp Glu Gly Asn Ala Ala Glu Gly Gly Phe Val Thr Glu Gly  
 1 5 10 15  
 Gly Lys Ser Glu Gly Met Lys Leu Trp Pro Leu Val Ile Phe Leu  
 20 25 30  
 Ser Tyr Phe Pro Gly Lys Pro Gly Glu Leu Thr Leu Phe Ser Val  
 35 40 45  
 Leu Pro Glu Leu Ser Gln Ser Leu Gly Leu Arg Glu Gln Glu Leu  
 50 55 60  
 Gln Val Val Arg Ala Ser Gly Lys Glu Ser Ser Gly Leu Val Leu  
 65 70 75  
 Leu Ser Ser Cys Pro Gln Thr Ala Ser Arg Leu Gln Lys Tyr Phe  
 80 85 90  
 Thr His Ala Arg Arg Ala Gln Arg Pro Thr Ala Thr Tyr Cys Ala  
 95 100 105  
 Val Thr Asp Gly Ile Pro Ala Ala Ser Glu Gly Lys Ile Gln Ala  
 110 115 120  
 Ala Leu Lys Leu Glu His Ile Asp Gly Val Asn Leu Thr Val Pro  
 125 130 135  
 Val Lys Ala Pro Ser Arg Lys Asp Ile Leu Glu Gly Val Lys Lys  
 140 145 150  
 Thr Leu Ser His Phe Arg Val Val Ala Thr Gly Ser Gly Cys Ala  
 155 160 165  
 Leu Val Gln Leu Gln Pro Leu Thr Val Phe Ser Ser Gln Leu Gln  
 170 175 180  
 Val His Met Val Leu Gln Leu Cys Pro Val Leu Gly Asp His Met  
 185 190 195  
 Tyr Ser Ala Arg Val Gly Thr Val Leu Gly Gln Arg Phe Leu Leu  
 200 205 210  
 Pro Ala Glu Asn Asn Lys Pro Gln Arg Gln Val Leu Asp Glu Ala  
 215 220 225  
 Leu Leu Arg Arg Leu His Leu Thr Pro Ser Gln Ala Ala Gln Leu  
 230 235 240  
 Pro Leu His Leu His Leu His Arg Leu Leu Leu Pro Gly Thr Arg  
 245 250 255  
 Ala Arg Asp Thr Pro Val Glu Leu Leu Ala Pro Leu Pro Pro Tyr  
 260 265 270  
 Phe Ser Arg Thr Leu Gln Cys Leu Gly Leu Arg Leu Gln  
 275 280

<210> 39  
 <211> 566  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 876751CD1

<400> 39  
 Met Asp Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu  
 1 5 10 15  
 Gln Ala Ala Ala Glu Phe Asp Gly Ser Arg Trp Pro Arg Gln Ile  
 20 25 30

Val	Ser	Ser	Ile	Gly	Leu	Cys	Arg	Tyr	Gly	Gly	Arg	Ile	Asp	Cys	
				35					40					45	
Cys	Trp	Gly	Trp	Ala	Arg	Gln	Ser	Trp	Gly	Gln	Cys	Gln	Pro	Val	
				50					55					60	
Cys	Gln	Pro	Arg	Cys	Lys	His	Gly	Glu	Cys	Ile	Gly	Pro	Asn	Lys	
				65					70					75	
Cys	Lys	Cys	His	Pro	Gly	Tyr	Ala	Gly	Lys	Thr	Cys	Asn	Gln	Asp	
				80					85					90	
Leu	Asn	Glu	Cys	Gly	Leu	Lys	Pro	Arg	Pro	Cys	Lys	His	Arg	Cys	
				95					100					105	
Met	Asn	Thr	Tyr	Gly	Ser	Tyr	Lys	Cys	Tyr	Cys	Leu	Asn	Gly	Tyr	
				110					115					120	
Met	Leu	Met	Pro	Asp	Gly	Ser	Cys	Ser	Ser	Ala	Leu	Thr	Cys	Ser	
				125					130					135	
Met	Ala	Asn	Cys	Gln	Tyr	Gly	Cys	Asp	Val	Val	Lys	Gly	Gln	Ile	
				140					145					150	
Arg	Cys	Gln	Cys	Pro	Ser	Pro	Gly	Leu	Gln	Leu	Ala	Pro	Asp	Gly	
				155					160					165	
Arg	Thr	Cys	Val	Asp	Val	Asp	Glu	Cys	Ala	Thr	Gly	Arg	Ala	Ser	
				170					175					180	
Cys	Pro	Arg	Phe	Arg	Gln	Cys	Val	Asn	Thr	Phe	Gly	Ser	Tyr	Ile	
				185					190					195	
Cys	Lys	Cys	His	Lys	Gly	Phe	Asp	Leu	Met	Tyr	Ile	Gly	Gly	Lys	
				200					205					210	
Tyr	Gln	Cys	His	Asp	Ile	Asp	Glu	Cys	Ser	Leu	Gly	Gln	Tyr	Gln	
				215					220					225	
Cys	Ser	Ser	Phe	Ala	Arg	Cys	Tyr	Asn	Val	Arg	Gly	Ser	Tyr	Lys	
				230					235					240	
Cys	Lys	Cys	Lys	Glu	Gly	Tyr	Gln	Gly	Asp	Gly	Leu	Thr	Cys	Val	
				245					250					255	
Tyr	Ile	Pro	Lys	Val	Met	Ile	Glu	Pro	Ser	Gly	Pro	Ile	His	Val	
				260					265					270	
Pro	Lys	Gly	Asn	Gly	Thr	Ile	Leu	Lys	Gly	Asp	Thr	Gly	Asn	Asn	
				275					280					285	
Asn	Trp	Ile	Pro	Asp	Val	Gly	Ser	Thr	Trp	Trp	Pro	Pro	Lys	Thr	
				290					295					300	
Pro	Tyr	Ile	Pro	Pro	Ile	Ile	Thr	Asn	Arg	Pro	Thr	Ser	Lys	Pro	
				305					310					315	
Thr	Thr	Arg	Pro	Thr	Pro	Lys	Pro	Thr	Pro	Ile	Pro	Thr	Pro	Pro	
				320					325					330	
Pro	Pro	Pro	Pro	Leu	Pro	Thr	Glu	Leu	Arg	Thr	Pro	Leu	Pro	Pro	
				335					340					345	
Thr	Thr	Pro	Glu	Arg	Pro	Thr	Thr	Gly	Leu	Thr	Thr	Ile	Ala	Pro	
				350					355					360	
Ala	Ala	Ser	Thr	Pro	Pro	Gly	Gly	Ile	Thr	Val	Asp	Asn	Arg	Val	
				365					370					375	
Gln	Thr	Asp	Pro	Gln	Lys	Pro	Arg	Gly	Asp	Val	Phe	Ile	Pro	Arg	
				380					385					390	
Gln	Pro	Ser	Asn	Asp	Leu	Phe	Glu	Ile	Phe	Glu	Ile	Glu	Arg	Gly	
				395					400					405	
Val	Ser	Ala	Asp	Asp	Glu	Ala	Lys	Asp	Asp	Pro	Gly	Val	Leu	Val	
				410					415					420	
His	Ser	Cys	Asn	Phe	Asp	His	Gly	Leu	Cys	Gly	Trp	Ile	Arg	Glu	
				425					430					435	
Lys	Asp	Asn	Asp	Leu	His	Trp	Glu	Pro	Ile	Arg	Asp	Pro	Ala	Gly	
				440					445					450	
Gly	Gln	Tyr	Leu	Thr	Val	Ser	Ala	Ala	Lys	Ala	Pro	Gly	Gly	Lys	
				455					460					465	
Ala	Ala	Arg	Leu	Val	Leu	Pro	Leu	Gly	Arg	Leu	Met	His	Ser	Gly	
				470					475					480	
Asp	Leu	Cys	Leu	Ser	Phe	Arg	His	Lys	Val	Thr	Gly	Leu	His	Ser	
				485					490					495	
Gly	Thr	Leu	Gln	Val	Phe	Val	Arg	Lys	His	Gly	Ala	His	Gly	Ala	

Ala Leu Trp Gly	Arg Asn Gly Gly His	Gly Trp Arg Gln Thr	Gln
500	505	510	
515	520	525	
Ile Thr Leu Arg	Gly Ala Asp Ile Lys	Ser Val Val Phe Lys	Gly
530	535	540	
Glu Lys Arg Arg	Gly His Thr Gly Glu	Ile Gly Leu Asp Asp	Val
545	550	555	
Ser Leu Lys Lys	Gly His Cys Ser Glu	Glu Arg	
560	565		

&lt;210&gt; 40

&lt;211&gt; 1093

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2512510CD1

&lt;400&gt; 40

Met Ala Arg Pro Val	Arg Gly Gly Leu	Gly Ala Pro Arg Arg	Ser
1	5	10	15
Pro Cys Leu Leu Leu	Leu Trp Leu Leu	Leu Leu Arg Leu Glu	Pro
20	25	30	
Val Thr Ala Ala Ala	Gly Pro Arg Ala	Pro Cys Ala Ala Ala	Cys
35	40	45	
Thr Cys Ala Gly Asp	Ser Leu Asp Cys	Gly Gly Arg Gly Leu	Ala
50	55	60	
Ala Leu Pro Gly Asp	Leu Pro Ser Trp	Thr Arg Ser Leu Asn	Leu
65	70	75	
Ser Tyr Asn Lys Leu	Ser Glu Ile Asp	Pro Ala Gly Phe Glu	Asp
80	85	90	
Leu Pro Asn Leu Gln	Glu Val Tyr Leu	Asn Asn Asn Glu Leu	Thr
95	100	105	
Ala Val Pro Ser Leu	Gly Ala Ala Ser	Ser His Val Val Ser	Leu
110	115	120	
Phe Leu Gln His Asn	Lys Ile Arg Ser	Val Glu Gly Ser Gln	Leu
125	130	135	
Lys Ala Tyr Leu Ser	Leu Glu Val Leu	Asp Leu Ser Leu Asn	Asn
140	145	150	
Ile Thr Glu Val Arg	Asn Thr Cys Phe	Pro His Gly Pro Pro	Ile
155	160	165	
Lys Glu Leu Asn Leu	Ala Gly Asn Arg	Ile Gly Thr Leu Glu	Leu
170	175	180	
Gly Ala Phe Asp Gly	Leu Ser Arg Ser	Leu Leu Thr Leu Arg	Leu
185	190	195	
Ser Lys Asn Arg Ile	Thr Gln Leu Pro	Val Arg Ala Phe Lys	Leu
200	205	210	
Pro Arg Leu Thr Gln	Leu Asp Leu Asn	Arg Asn Arg Ile Arg	Leu
215	220	225	
Ile Glu Gly Leu Thr	Phe Gln Gly Leu	Asn Ser Leu Glu Val	Leu
230	235	240	
Lys Leu Gln Arg Asn	Asn Ile Ser Lys	Leu Thr Asp Gly Ala	Phe
245	250	255	
Trp Gly Leu Ser Lys	Met His Val Leu	His Leu Glu Tyr Asn	Ser
260	265	270	
Leu Val Glu Val Asn	Ser Gly Ser Leu	Tyr Gly Leu Thr Ala	Leu
275	280	285	
His Gln Leu His Leu	Ser Asn Asn Ser	Ile Ala Arg Ile His	Arg
290	295	300	
Lys Gly Trp Ser Phe	Cys Gln Lys Leu	His Glu Leu Val Leu	Ser
305	310	315	
Phe Asn Asn Leu Thr	Arg Leu Asp Glu	Glu Ser Leu Ala Glu	Leu



	320		325		330
Ser Ser Leu Ser	Val Leu Arg Leu Ser	His Asn Ser Ile Ser	His		
	335		340		345
Ile Ala Glu Gly	Ala Phe Lys Gly Leu	Arg Ser Leu Arg Val	Leu		
	350		355		360
Asp Leu Asp His	Asn Glu Ile Ser Gly	Thr Ile Glu Asp Thr	Ser		
	365		370		375
Gly Ala Phe Ser	Gly Leu Asp Ser Leu	Ser Lys Leu Thr Leu	Phe		
	380		385		390
Gly Asn Lys Ile	Lys Ser Val Ala Lys	Arg Ala Phe Ser Gly	Leu		
	395		400		405
Glu Gly Leu Glu	His Leu Asn Leu Gly	Gly Asn Ala Ile Arg	Ser		
	410		415		420
Val Gln Phe Asp	Ala Phe Val Lys Met	Lys Asn Leu Lys Glu	Leu		
	425		430		435
His Ile Ser Ser	Asp Ser Phe Leu Cys	Asp Cys Gln Leu Lys	Trp		
	440		445		450
Leu Pro Pro Trp	Leu Ile Gly Arg Met	Leu Gln Ala Phe Val	Thr		
	455		460		465
Ala Thr Cys Ala	His Pro Glu Ser Leu	Lys Gly Gln Ser Ile	Phe		
	470		475		480
Ser Val Pro Pro	Glu Ser Phe Val Cys	Asp Asp Phe Leu Lys	Pro		
	485		490		495
Gln Ile Ile Thr	Gln Pro Glu Thr Thr	Met Ala Met Val Gly	Lys		
	500		505		510
Asp Ile Arg Phe	Thr Cys Ser Ala Ala	Ser Ser Ser Ser	Pro		
	515		520		525
Met Thr Phe Ala	Trp Lys Lys Asp Asn	Glu Val Leu Thr Asn	Ala		
	530		535		540
Asp Met Glu Asn	Phe Val His Val His	Ala Gln Asp Gly Glu	Val		
	545		550		555
Met Glu Tyr Thr	Thr Ile Leu His Leu	Arg Gln Val Thr Phe	Gly		
	560		565		570
His Glu Gly Arg	Tyr Gln Cys Val Ile	Thr Asn His Phe Gly	Ser		
	575		580		585
Thr Tyr Ser His	Lys Ala Arg Leu Thr	Val Asn Val Leu Pro	Ser		
	590		595		600
Phe Thr Lys Thr	Pro His Asp Ile Thr	Ile Arg Thr Thr Thr	Met		
	605		610		615
Ala Arg Leu Glu	Cys Ala Ala Thr Gly	His Pro Asn Pro Gln	Ile		
	620		625		630
Ala Trp Gln Lys	Asp Gly Gly Thr Asp	Phe Pro Ala Ala Arg	Glu		
	635		640		645
Arg Arg Met His	Val Met Pro Asp Asp	Asp Val Phe Phe Ile	Thr		
	650		655		660
Asp Val Lys Ile	Asp Asp Ala Gly Val	Tyr Ser Cys Thr Ala	Gln		
	665		670		675
Asn Ser Ala Gly	Ser Ile Ser Ala Asn	Ala Thr Leu Thr Val	Leu		
	680		685		690
Glu Thr Pro Ser	Leu Val Val Pro Leu	Glu Asp Arg Val Val	Ser		
	695		700		705
Val Gly Glu Thr	Val Ala Leu Gln Cys	Lys Ala Thr Gly Asn	Pro		
	710		715		720
Pro Pro Arg Ile	Thr Trp Phe Lys Gly	Asp Arg Pro Leu Ser	Leu		
	725		730		735
Thr Glu Arg His	His Leu Thr Pro Asp	Asn Gln Leu Leu Val	Val		
	740		745		750
Gln Asn Val Val	Ala Glu Asp Ala Gly	Arg Tyr Thr Cys Glu	Met		
	755		760		765
Ser Asn Thr Leu	Gly Thr Glu Arg Ala	His Ser Gln Leu Ser	Val		
	770		775		780
Leu Pro Ala Ala	Gly Cys Arg Lys Asp	Gly Thr Thr Val Gly	Ile		
	785		790		795

```

Phe Thr Ile Ala Val Val Ser Ser Ile Val Leu Thr Ser Leu Val
      800                      805                      810
Trp Val Cys Ile Ile Tyr Gln Thr Arg Lys Lys Ser Glu Glu Tyr
      815                      820                      825
Ser Val Thr Asn Thr Asp Glu Thr Val Val Pro Pro Asp Val Pro
      830                      835                      840
Ser Tyr Leu Ser Ser Gln Gly Thr Leu Ser Asp Arg Gln Glu Thr
      845                      850                      855
Val Val Arg Thr Glu Gly Gly Pro Gln Ala Asn Gly His Ile Glu
      860                      865                      870
Ser Asn Gly Val Cys Pro Arg Asp Ala Ser His Phe Pro Glu Pro
      875                      880                      885
Asp Thr His Ser Val Ala Cys Arg Gln Pro Lys Leu Cys Ala Gly
      890                      895                      900
Ser Ala Tyr His Lys Glu Pro Trp Lys Ala Met Glu Lys Ala Glu
      905                      910                      915
Gly Thr Pro Gly Pro His Lys Met Glu His Gly Gly Arg Val Val
      920                      925                      930
Cys Ser Asp Cys Asn Thr Glu Val Asp Cys Tyr Ser Arg Gly Gln
      935                      940                      945
Ala Phe His Pro Gln Pro Val Ser Arg Asp Ser Ala Gln Pro Ser
      950                      955                      960
Ala Pro Asn Gly Pro Glu Pro Gly Gly Ser Asp Gln Glu His Ser
      965                      970                      975
Pro His His Gln Cys Ser Arg Thr Ala Ala Gly Ser Cys Pro Glu
      980                      985                      990
Cys Gln Gly Ser Leu Tyr Pro Ser Asn His Asp Arg Met Leu Thr
      995                      1000                     1005
Ala Val Lys Lys Lys Pro Met Ala Ser Leu Asp Gly Lys Gly Asp
      1010                     1015                     1020
Ser Ser Trp Thr Thr Ala Arg Leu Tyr His Pro Asp Ser Thr Glu
      1025                     1030                     1035
Leu Gln Pro Ala Ser Ser Leu Thr Ser Gly Ser Pro Glu Arg Ala
      1040                     1045                     1050
Glu Ala Gln Tyr Leu Leu Val Ser Asn Gly His Leu Pro Lys Ala
      1055                     1060                     1065
Cys Asp Ala Ser Pro Glu Ser Thr Pro Leu Thr Gly Gln Leu Pro
      1070                     1075                     1080
Gly Lys Gln Arg Val Pro Leu Leu Leu Ala Pro Lys Ser
      1085                     1090

```

<210> 41  
 <211> 915  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7486326CD1

```

<400> 41
Met Pro Ser Leu Pro Ala Pro Pro Ala Pro Leu Leu Leu Leu Gly
  1          5          10          15
Leu Leu Leu Leu Gly Ser Arg Pro Ala Arg Gly Ala Gly Pro Glu
  20          25          30
Pro Pro Val Leu Pro Ile Arg Ser Glu Lys Glu Pro Leu Pro Val
  35          40          45
Arg Gly Ala Ala Gly Cys Thr Phe Gly Gly Lys Val Tyr Ala Leu
  50          55          60
Asp Glu Thr Trp His Pro Asp Leu Gly Glu Pro Phe Gly Val Met
  65          70          75
Arg Cys Val Leu Cys Ala Cys Glu Ala Pro Gln Trp Gly Arg Arg
  80          85          90

```

Thr	Arg	Gly	Pro	Gly	Arg	Val	Ser	Cys	Lys	Asn	Ile	Lys	Pro	Glu	
				95					100					105	
Cys	Pro	Thr	Pro	Ala	Cys	Gly	Gln	Pro	Arg	Gln	Leu	Pro	Gly	His	
				110					115					120	
Cys	Cys	Gln	Thr	Cys	Pro	Gln	Glu	Arg	Ser	Ser	Ser	Glu	Arg	Gln	
				125					130					135	
Pro	Ser	Gly	Leu	Ser	Phe	Glu	Tyr	Pro	Arg	Asp	Pro	Glu	His	Arg	
				140					145					150	
Ser	Tyr	Ser	Asp	Arg	Gly	Glu	Pro	Gly	Ala	Glu	Glu	Arg	Ala	Arg	
				155					160					165	
Gly	Asp	Gly	His	Thr	Asp	Phe	Val	Ala	Leu	Leu	Thr	Gly	Pro	Arg	
				170					175					180	
Ser	Gln	Ala	Val	Ala	Arg	Ala	Arg	Val	Ser	Leu	Leu	Arg	Ser	Ser	
				185					190					195	
Leu	Arg	Phe	Ser	Ile	Ser	Tyr	Arg	Arg	Leu	Asp	Arg	Pro	Thr	Arg	
				200					205					210	
Ile	Arg	Phe	Ser	Asp	Ser	Asn	Gly	Ser	Val	Leu	Phe	Glu	His	Pro	
				215					220					225	
Ala	Ala	Pro	Thr	Gln	Asp	Gly	Leu	Val	Cys	Gly	Val	Trp	Arg	Ala	
				230					235					240	
Val	Pro	Arg	Leu	Ser	Leu	Arg	Leu	Leu	Arg	Ala	Glu	Gln	Leu	His	
				245					250					255	
Val	Ala	Leu	Val	Thr	Leu	Thr	His	Pro	Ser	Gly	Glu	Val	Trp	Gly	
				260					265					270	
Pro	Leu	Ile	Arg	His	Arg	Ala	Leu	Ala	Ala	Glu	Thr	Phe	Ser	Ala	
				275					280					285	
Ile	Leu	Thr	Leu	Glu	Gly	Pro	Pro	Gln	Gln	Gly	Val	Gly	Gly	Ile	
				290					295					300	
Thr	Leu	Leu	Thr	Leu	Ser	Asp	Thr	Glu	Asp	Ser	Leu	His	Phe	Leu	
				305					310					315	
Leu	Leu	Phe	Arg	Gly	Leu	Leu	Glu	Pro	Arg	Ser	Gly	Gly	Leu	Thr	
				320					325					330	
Gln	Val	Pro	Leu	Arg	Leu	Gln	Ile	Leu	His	Gln	Gly	Gln	Leu	Leu	
				335					340					345	
Arg	Glu	Leu	Gln	Ala	Asn	Val	Ser	Ala	Gln	Glu	Pro	Gly	Phe	Ala	
				350					355					360	
Glu	Val	Leu	Pro	Asn	Leu	Thr	Val	Gln	Glu	Met	Asp	Trp	Leu	Val	
				365					370					375	
Leu	Gly	Glu	Leu	Gln	Met	Ala	Leu	Glu	Trp	Ala	Gly	Arg	Pro	Gly	
				380					385					390	
Leu	Arg	Ile	Ser	Gly	His	Ile	Ala	Ala	Arg	Lys	Ser	Cys	Asp	Val	
				395					400					405	
Leu	Gln	Ser	Val	Leu	Cys	Gly	Ala	Asp	Ala	Leu	Ile	Pro	Val	Gln	
				410					415					420	
Thr	Gly	Ala	Ala	Gly	Ser	Ala	Ser	Leu	Thr	Leu	Leu	Gly	Asn	Gly	
				425					430					435	
Ser	Leu	Ile	Tyr	Gln	Ala	Val	Gly	Ile	Cys	Pro	Gly	Leu	Gly	Ala	
				440					445					450	
Arg	Gly	Ala	His	Met	Leu	Leu	Gln	Asn	Glu	Leu	Phe	Leu	Asn	Val	
				455					460					465	
Gly	Thr	Lys	Asp	Phe	Pro	Asp	Gly	Glu	Leu	Arg	Gly	His	Val	Ala	
				470					475					480	
Ala	Leu	Pro	Tyr	Cys	Gly	His	Ser	Ala	Arg	His	Asp	Thr	Leu	Pro	
				485					490					495	
Val	Pro	Leu	Ala	Gly	Ala	Leu	Val	Leu	Pro	Pro	Val	Lys	Ser	Gln	
				500					505					510	
Ala	Ala	Gly	His	Ala	Trp	Leu	Ser	Leu	Asp	Thr	His	Cys	His	Leu	
				515					520					525	
His	Tyr	Glu	Val	Leu	Leu	Ala	Gly	Leu	Gly	Gly	Ser	Glu	Gln	Gly	
				530					535					540	
Thr	Val	Thr	Ala	His	Leu	Leu	Gly	Pro	Pro	Gly	Thr	Pro	Gly	Pro	
				545					550					555	
Arg	Arg	Leu	Leu	Lys	Gly	Phe	Tyr	Gly	Ser	Glu	Ala	Gln	Gly	Val	

Val Lys Asp Leu	560	Glu Pro Glu Leu Leu	565	Arg His Leu Ala Lys	570
	575		580		585
Met Ala Ser Leu	590	Leu Ile Thr Thr Lys	595	Gly Ser Pro Arg Gly	600
	605		610		615
Leu Arg Gly Gln	620	Val His Ile Ala Asn	625	Gln Cys Glu Val Gly	630
	635		640		645
Leu Arg Leu Glu	650	Ala Ala Gly Ala Glu	655	Gly Val Arg Ala Leu	660
	665		670		675
Ala Pro Asp Pro	680	Ala Ser Ala Ala Pro	685	Pro Val Val Pro Gly	690
	695		700		705
Pro Ala Leu Ala	710	Pro Ala Lys Pro Gly	715	Gly Pro Gly Arg Pro	720
	725		730		735
Asp Pro Asn Thr	740	Cys Phe Phe Glu Gly	745	Gln Gln Arg Pro His	750
	755		760		765
Ala Arg Trp Ala	770	Pro Asn Tyr Asp Pro	775	Leu Cys Ser Leu Cys	780
	785		790		795
Cys Gln Arg Arg	800	Thr Val Ile Cys Asp	805	Pro Val Val Cys Pro	810
	815		820		825
Pro Ser Cys Pro	830	His Pro Val Gln Ala	835	Pro Asp Gln Cys Cys	840
	845		850		855
Val Cys Pro Glu	860	Lys Gln Asp Val Arg	865	Asp Leu Pro Gly Leu	870
	875		880		885
Arg Ser Arg Asp	890	Pro Gly Glu Gly Cys	895	Tyr Phe Asp Gly Asp	900
	905		910		915
Ser Trp Arg Ala		Ala Gly Thr Arg Trp		His Pro Val Val Pro	
Phe Gly Leu Ile		Lys Cys Ala Val Cys		Thr Cys Lys Gly Gly	
Gly Glu Val His		Cys Glu Lys Val Gln		Cys Pro Arg Leu Ala	
Ala Gln Pro Val		Arg Val Asn Pro Thr		Asp Cys Cys Lys Gln	
Pro Val Gly Ser		Gly Ala His Pro Gln		Leu Gly Asp Pro Met	
Ala Asp Gly Pro		Arg Gly Cys Arg Phe		Ala Gly Gln Trp Phe	
Glu Ser Gln Ser		Trp His Pro Ser Val		Pro Pro Phe Gly Glu	
Ser Cys Ile Thr		Cys Arg Cys Gly Ala		Gly Val Pro His Cys	
Arg Asp Asp Cys		Ser Leu Pro Leu Ser		Cys Gly Ser Gly Lys	
Ser Arg Cys Cys		Ser Arg Cys Thr Ala		His Arg Arg Pro Ala	
Glu Thr Arg Thr		Asp Pro Glu Leu Glu		Lys Glu Ala Glu Gly	

&lt;210&gt; 42

&lt;211&gt; 113

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1221545CD1

&lt;400&gt; 42

Met Ala Trp Thr	Leu Ala Cys Val Cys	Val Leu Gly Ser Ile	Leu
1	5	10	15
Val Leu Asp Ser	Gly Met Cys Val Arg	Ala Gly Glu Cys Leu	Asp
	20	25	30

```

Gly Asp Val Val Ser Leu Leu His Phe Trp His Ser Val Thr Thr
      35              40              45
Gln Glu Asn Gln Ile Glu Asn Leu Glu Ser Val Leu Gln Trp Ile
      50              55              60
Glu Thr Gly Leu Gln Ser Leu Arg Lys Lys Ser Lys Gln Asn Thr
      65              70              75
Gln Glu Phe Arg Glu Asn Ile Phe Leu Pro Lys Asn Asn Phe Ser
      80              85              90
Phe Met Leu Phe Leu Ile Trp Val Asn Thr Pro Met Glu Lys Ile
      95              100             105
Asp Arg Leu Val Lys Ser Ser Ile
      110

```

<210> 43  
 <211> 91  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 124737CD1

```

<400> 43
Met Gly Lys Gly Arg Trp Ala Thr Val Gly Val Ser Pro Cys Leu
  1      5      10      15
Pro Pro Leu Trp Ala Ala Ala Gly Ala His Ala Ser Lys Ser Ser
      20      25      30
Leu Arg Glu Arg Glu Leu Arg Cys Leu Tyr Pro Ser Ser Val Arg
      35      40      45
His Trp Leu Asn Val His Thr Pro Gly Ser Pro Pro Leu Ile Leu
      50      55      60
Met Met Ser His Gly Pro His Phe Thr Ser Glu Leu Trp Val His
      65      70      75
Gly Glu His Gln Ser His Pro Gly Ser Val Pro Gln Leu Ser Leu
      80      85      90
Thr

```

<210> 44  
 <211> 83  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1510784CD1

```

<400> 44
Met Arg Met Phe Pro Leu Pro Leu Pro Val Cys Leu Pro Leu Gly
  1      5      10      15
Val His Leu Gln Ser Thr Ser Pro Pro Phe Pro Ala Ser His Thr
      20      25      30
Gln Val Ser Leu Ser Asp Ser His Thr Cys Leu Thr Ala Ser Pro
      35      40      45
Ala Lys Val Leu Phe Lys Cys Leu Phe Ser Val Cys Leu Cys His
      50      55      60
Ser Gln Cys Asp His Ser Cys Ser Ala Val Ser Gln Gln Glu Asp
      65      70      75
Arg Cys Arg Ser Ser Ser Cys Ser
      80

```

<210> 45  
 <211> 128

<212> PRT  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1901257CD1

<400> 45  
Met Pro Tyr Ala Leu His Met Ser Phe Gln Arg Leu Trp Val Trp  
1 5 10 15  
Ile Leu Leu Pro Thr Val Ala Asn Ile Ala Leu Ser Ser Ser Arg  
20 25 30  
Thr Gly Arg Ser Lys Glu His Thr Gln Asp Asp Ala Thr Ala Tyr  
35 40 45  
Met Leu Ser Arg His Leu His Ala Leu Ser Ala Pro Thr Cys Ser  
50 55 60  
Leu Gly Ser Leu His Ala Leu Ser Ala Ala Tyr Thr Leu Ser Trp  
65 70 75  
His Val Gln Gln Val Leu Gln Pro Cys Pro Gly Gly Leu Gly Leu  
80 85 90  
Arg Gly Leu Ser Leu Ser Trp Val Leu Asp Leu Pro Pro His Phe  
95 100 105  
His His Cys Asn Phe Cys Phe Thr Cys Trp Lys Gly Ala Ser Tyr  
110 115 120  
Asn Met Pro Leu Lys Glu Lys Asp  
125

<210> 46  
<211> 84  
<212> PRT  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2044370CD1

<400> 46  
Met Ala Leu Leu Trp Trp Ile Ser Thr Val Ala Ile Leu Leu Phe  
1 5 10 15  
Thr Ser Thr Ile Leu Gly Thr Tyr Val Glu Ala Gly Ala Ala Lys  
20 25 30  
Ser Asn Glu Glu Glu Ile Val Asn Lys Ser Glu Phe Gly Arg Phe  
35 40 45  
Pro Arg Gly Ser Arg Lys Asp Ala Ser Gly Cys His Lys Pro Gly  
50 55 60  
Tyr Pro Val Pro Pro His Ser Arg Cys Pro Pro Pro Pro His Val  
65 70 75  
Gln Arg Pro Arg Pro Ile Leu His Ala  
80

<210> 47  
<211> 109  
<212> PRT  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2820933CD1

<400> 47  
Met Gly Trp Pro Pro Pro Pro Gly Ser Ser Phe Cys Leu Cys Phe  
1 5 10 15  
Ile His Gly Ala Phe Ser Ser Phe Ser Pro His Pro Pro Ser His

	20		25		30
Glu Cys Ser Ser Arg	Cys Cys Ser Leu	Cys Leu Ala Arg Phe	Leu		
	35		40		45
Ala Ser Pro Leu Pro	Trp Ser Asn Ser	Glu Ser Ser Ser Thr	Leu		
	50		55		60
Tyr Leu Lys Ser Arg	Leu Ala Gly Ser	Leu Ser Gly Ser Ala	His		
	65		70		75
Cys Ser Pro Thr Ser	Leu Pro Phe Ser	Leu Gly Thr Leu Ile	Thr		
	80		85		90
Pro Glu Thr Val Asp	Ser Ser Pro Lys	Tyr Ser Phe Trp Leu	Ile		
	95		100		105
Val Gly Ala Gln					

<210> 48  
 <211> 159  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2902793CD1

<400> 48	
Met Trp Ser Val Ser	Ser Trp Ala Leu Cys
1	5
	10
His Val Leu Ser Leu	Ser Cys Ala Gln Cys
	20
	25
Phe Leu Ile Pro Pro	Pro Ala Leu Pro Ala
	35
	40
Leu Arg Asn Glu Glu	Ala Met Glu Gly Ala
	50
	55
Cys Glu Leu Ser Lys	Ala Ala Pro Val Glu
	65
	70
Glu Ala Leu Arg Asp	Gly Asp Lys Tyr Ser
	80
	85
Ala Val Cys Glu Leu	Gln Ile His Gly Leu
	95
	100
Gly Val Tyr Ser Cys	Val Cys Gly Gln Glu
	110
	115
Leu Thr Val Arg Gly	Lys Asp Pro Met Trp
	125
	130
Ala Trp Cys Ile His	Leu Ser Val Ser Pro
	140
	145
Cys Gly Thr Ser Pro	Val Glu Thr Leu
	155

<210> 49  
 <211> 242  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7486536CD1

<400> 49	
Met Pro Arg Gly Phe	Thr Trp Leu Arg Tyr
1	5
	10
Gly Val Ala Leu Gly	Asn Glu Pro Leu Glu
	20
	25
Gln Asn Glu Glu Cys	Thr Val Thr Gly Phe
	35
	40
	45

Gln Tyr Arg Ser Arg Leu Gln Tyr Met Lys His Tyr Phe Pro Ile  
 50 55 60  
 Asn Tyr Lys Ile Ser Val Pro Tyr Glu Gly Val Phe Arg Ile Ala  
 65 70 75  
 Asn Val Thr Arg Leu Gln Arg Ala Gln Val Ser Glu Arg Glu Leu  
 80 85 90  
 Arg Tyr Leu Trp Val Leu Val Ser Leu Ser Ala Thr Glu Ser Val  
 95 100 105  
 Gln Asp Val Leu Leu Glu Gly His Pro Ser Trp Lys Tyr Leu Gln  
 110 115 120  
 Glu Val Glu Thr Leu Leu Leu Asn Val Gln Gln Gly Leu Thr Asp  
 125 130 135  
 Val Glu Val Ser Pro Lys Val Glu Ser Val Leu Ser Leu Leu Asn  
 140 145 150  
 Ala Pro Gly Pro Asn Leu Lys Leu Val Arg Pro Lys Ala Leu Leu  
 155 160 165  
 Asp Asn Cys Phe Arg Val Met Glu Leu Leu Tyr Cys Ser Cys Cys  
 170 175 180  
 Lys Gln Ser Ser Val Leu Asn Trp Gln Asp Cys Glu Val Pro Ser  
 185 190 195  
 Pro Gln Ser Cys Ser Pro Glu Pro Ser Leu Gln Tyr Ala Ala Thr  
 200 205 210  
 Gln Leu Tyr Pro Pro Pro Pro Trp Ser Pro Ser Ser Pro Pro His  
 215 220 225  
 Ser Thr Gly Ser Val Arg Pro Val Arg Ala Gln Gly Glu Gly Leu  
 230 235 240  
 Leu Pro

&lt;210&gt; 50

&lt;211&gt; 542

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 8137305CD1

&lt;400&gt; 50

Met Pro Arg Arg Gly Leu Ile Leu His Thr Arg Thr His Trp Leu  
 1 5 10 15  
 Leu Leu Gly Leu Ala Leu Leu Cys Ser Leu Val Leu Phe Met Tyr  
 20 25 30  
 Leu Leu Glu Cys Ala Pro Gln Thr Asp Gly Asn Ala Ser Leu Pro  
 35 40 45  
 Gly Val Val Gly Glu Asn Tyr Gly Lys Glu Tyr Tyr Gln Ala Leu  
 50 55 60  
 Leu Gln Glu Gln Glu Glu His Tyr Gln Thr Arg Ala Thr Ser Leu  
 65 70 75  
 Lys Arg Gln Ile Ala Gln Leu Lys Gln Glu Leu Gln Glu Met Ser  
 80 85 90  
 Glu Lys Met Arg Ser Leu Gln Glu Arg Arg Asn Val Gly Ala Asn  
 95 100 105  
 Gly Ile Gly Tyr Gln Ser Asn Lys Glu Gln Ala Pro Ser Asp Leu  
 110 115 120  
 Leu Glu Phe Leu His Ser Gln Ile Asp Lys Ala Glu Val Ser Ile  
 125 130 135  
 Gly Ala Lys Leu Pro Ser Glu Tyr Gly Val Ile Pro Phe Glu Ser  
 140 145 150  
 Phe Thr Leu Met Lys Val Phe Gln Leu Glu Met Gly Leu Thr Arg  
 155 160 165  
 His Pro Glu Glu Lys Pro Val Arg Lys Asp Lys Arg Asp Glu Leu  
 170 175 180



Val	Glu	Val	Ile	Glu	Ala	Gly	Leu	Glu	Val	Ile	Asn	Asn	Pro	Asp	
				185					190						195
Glu	Asp	Asp	Glu	Gln	Glu	Asp	Glu	Glu	Gly	Pro	Leu	Gly	Glu	Lys	
				200					205						210
Leu	Ile	Phe	Asn	Glu	Asn	Asp	Phe	Val	Glu	Gly	Tyr	Tyr	Arg	Thr	
				215					220						225
Glu	Arg	Asp	Lys	Gly	Thr	Gln	Tyr	Glu	Leu	Phe	Phe	Lys	Lys	Ala	
				230					235						240
Asp	Leu	Thr	Glu	Tyr	Arg	His	Val	Thr	Leu	Phe	Arg	Pro	Phe	Gly	
				245					250						255
Pro	Leu	Met	Lys	Val	Lys	Ser	Glu	Met	Ile	Asp	Ile	Thr	Arg	Ser	
				260					265						270
Ile	Ile	Asn	Ile	Ile	Val	Pro	Leu	Ala	Glu	Arg	Thr	Glu	Ala	Phe	
				275					280						285
Val	Gln	Phe	Met	Gln	Asn	Phe	Arg	Asp	Val	Cys	Ile	His	Gln	Asp	
				290					295						300
Lys	Lys	Ile	His	Leu	Thr	Val	Val	Tyr	Phe	Gly	Lys	Glu	Gly	Leu	
				305					310						315
Ser	Lys	Val	Lys	Ser	Ile	Leu	Glu	Ser	Val	Thr	Ser	Glu	Ser	Asn	
				320					325						330
Phe	His	Asn	Tyr	Thr	Leu	Val	Ser	Leu	Asn	Glu	Glu	Phe	Asn	Arg	
				335					340						345
Gly	Arg	Gly	Leu	Asn	Val	Gly	Ala	Arg	Ala	Trp	Asp	Lys	Gly	Glu	
				350					355						360
Val	Leu	Met	Phe	Phe	Cys	Asp	Val	Asp	Ile	Tyr	Phe	Ser	Ala	Glu	
				365					370						375
Phe	Leu	Asn	Ser	Cys	Arg	Leu	Asn	Ala	Glu	Pro	Gly	Lys	Lys	Val	
				380					385						390
Phe	Tyr	Pro	Val	Val	Phe	Ser	Leu	Tyr	Asn	Pro	Ala	Ile	Val	Tyr	
				395					400						405
Ala	Asn	Gln	Glu	Val	Pro	Pro	Pro	Val	Glu	Gln	Gln	Leu	Val	His	
				410					415						420
Lys	Lys	Asp	Ser	Gly	Phe	Trp	Arg	Asp	Phe	Gly	Phe	Gly	Met	Thr	
				425					430						435
Cys	Gln	Tyr	Arg	Ser	Asp	Phe	Leu	Thr	Ile	Gly	Gly	Phe	Asp	Met	
				440					445						450
Glu	Val	Lys	Gly	Trp	Gly	Gly	Glu	Asp	Val	His	Leu	Tyr	Arg	Lys	
				455					460						465
Tyr	Leu	His	Gly	Asp	Leu	Ile	Val	Ile	Arg	Thr	Pro	Val	Pro	Gly	
				470					475						480
Leu	Phe	His	Leu	Trp	His	Glu	Lys	Arg	Cys	Ala	Asp	Glu	Leu	Thr	
				485					490						495
Pro	Glu	Gln	Tyr	Arg	Met	Cys	Ile	Gln	Ser	Lys	Ala	Met	Asn	Glu	
				500					505						510
Ala	Ser	His	Ser	His	Leu	Gly	Met	Leu	Val	Phe	Arg	Glu	Glu	Ile	
				515					520						525
Glu	Thr	His	Leu	His	Lys	Gln	Ala	Tyr	Arg	Thr	Asn	Ser	Glu	Ala	
				530					535						540
Val	Gly														

&lt;210&gt; 51

&lt;211&gt; 105

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3793128CD1

&lt;400&gt; 51

Met	Ser	His	Leu	Leu	Ala	Pro	Asn	Leu	Phe	Phe	Val	Leu	Leu	Asn	
1				5					10					15	

Leu	Val	Thr	Ser	Leu	Leu	Arg	Leu	Ile	Gly	Val	Gln	His	Lys	Ser	
				20					25					30	
Phe	Arg	Ser	Tyr	Leu	Ala	Thr	Pro	Arg	Pro	Phe	Ala	Phe	Leu	Lys	
				35					40					45	
Glu	Glu	Ile	Ile	Gly	Thr	Leu	Leu	Leu	Asn	Gly	Thr	Tyr	Thr	Ala	
				50					55					60	
Val	Val	Cys	Tyr	Phe	Tyr	Lys	Gly	Ser	Gln	Ala	Phe	Thr	Cys	Phe	
				65					70					75	
Pro	His	Phe	Asn	Leu	Pro	Cys	Ala	Cys	Arg	Val	Ile	Val	Arg	Asp	
				80					85					90	
Phe	Arg	Asn	Pro	Arg	Ser	Trp	Val	Pro	Phe	Trp	Thr	Leu	Cys	His	
				95					100					105	

<210> 52  
 <211> 102  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4001243CD1

<400>	52															
Met	Arg	Leu	Arg	His	Arg	Gln	Arg	Ala	Leu	Pro	Thr	Thr	Leu	Ala		
1				5					10					15		
Thr	Ala	Ser	Lys	Pro	Leu	Phe	Met	Pro	Gly	Thr	Ala	Pro	Lys	Asp		
				20					25					30		
Leu	Ala	His	Ala	Trp	Asp	Arg	Pro	Gln	Gly	Pro	His	Trp	Leu	Gln		
				35					40					45		
Ser	Ala	Ala	Gly	Arg	Val	Val	Gly	Glu	Gly	Met	Asp	Thr	Pro	Trp		
				50					55					60		
Ala	Gly	Ala	Gly	Arg	Thr	Arg	Pro	Ile	Ile	Gly	His	Leu	Val	Ala		
				65					70					75		
Met	Ala	Thr	Thr	Gln	Gly	Cys	Leu	Arg	Leu	Lys	Ile	Cys	Gly	Leu		
				80					85					90		
Gln	Gly	Ala	Pro	Ala	Leu	Ala	Leu	Ala	Glu	Ser	Gln					
				95					100							

<210> 53  
 <211> 129  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 6986717CD1

<400>	53															
Met	Val	Ile	Pro	Gly	Leu	Thr	Thr	Leu	Leu	Ile	Lys	Thr	Thr	Phe		
1				5					10					15		
Trp	Gly	Phe	Arg	Phe	Gly	Glu	Leu	Gly	Met	Gly	Arg	Gly	Ser	Thr		
				20					25					30		
Ser	Ser	Arg	Cys	Leu	Val	Ser	Pro	Ser	Phe	Ser	Leu	Leu	His	Val		
				35					40					45		
Gly	Gly	Arg	Leu	Asp	Gln	Leu	Ala	Cys	Thr	Leu	Pro	Lys	Glu	Leu		
				50					55					60		
Arg	Gly	Lys	Asp	Met	Arg	Met	Val	Pro	Met	Glu	Met	Phe	Asn	Tyr		
				65					70					75		
Cys	Ser	Gln	Leu	Glu	Asp	Glu	Asn	Ser	Ser	Ala	Gly	Leu	Asp	Ile		
				80					85					90		
Leu	Gly	His	Pro	Ala	Pro	Arg	Pro	Val	Gln	Ser	Leu	Leu	Ser	Pro		
				95					100					105		

Ser Pro Gly Leu Ser Arg Ser Arg Ser Pro Ala Gln Pro Ala His  
 110 115 120  
 Arg Ser Arg Gly Thr Gly Arg Arg Ala  
 125

<210> 54  
 <211> 1070  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7503512CD1

<400> 54  
 Met Ala Arg Pro Val Arg Gly Gly Leu Gly Ala Pro Arg Arg Ser  
 1 5 10 15  
 Pro Cys Leu Leu Leu Leu Trp Leu Leu Leu Leu Arg Leu Glu Pro  
 20 25 30  
 Val Thr Ala Ala Ala Gly Pro Arg Ala Pro Cys Ala Ala Ala Cys  
 35 40 45  
 Thr Cys Ala Gly Asp Ser Leu Asp Cys Gly Gly Arg Gly Leu Ala  
 50 55 60  
 Ala Leu Pro Gly Asp Leu Pro Ser Trp Thr Arg Ser Leu Asn Leu  
 65 70 75  
 Ser Tyr Asn Lys Leu Ser Glu Ile Asp Pro Ala Gly Phe Glu Asp  
 80 85 90  
 Leu Pro Asn Leu Gln Glu Val Tyr Leu Asn Asn Asn Glu Leu Thr  
 95 100 105  
 Ala Val Pro Ser Leu Gly Ala Ala Ser Ser His Val Val Ser Leu  
 110 115 120  
 Phe Leu Gln His Asn Lys Ile Arg Ser Val Glu Gly Ser Gln Leu  
 125 130 135  
 Lys Ala Tyr Leu Ser Leu Glu Val Leu Asp Leu Ser Leu Asn Asn  
 140 145 150  
 Ile Thr Glu Val Arg Asn Thr Cys Phe Pro His Gly Pro Pro Ile  
 155 160 165  
 Lys Glu Leu Asn Leu Ala Gly Asn Arg Ile Gly Thr Leu Glu Leu  
 170 175 180  
 Gly Ala Phe Asp Gly Leu Ser Arg Ser Leu Leu Thr Leu Arg Leu  
 185 190 195  
 Ser Lys Asn Arg Ile Arg Leu Ile Glu Gly Leu Thr Phe Gln Gly  
 200 205 210  
 Leu Asn Ser Leu Glu Val Leu Lys Leu Gln Arg Asn Asn Ile Ser  
 215 220 225  
 Lys Leu Thr Asp Gly Ala Phe Trp Gly Leu Ser Lys Met His Val  
 230 235 240  
 Leu His Leu Glu Tyr Asn Ser Leu Val Glu Val Asn Ser Gly Ser  
 245 250 255  
 Leu Tyr Gly Leu Thr Ala Leu His Gln Leu His Leu Ser Asn Asn  
 260 265 270  
 Ser Ile Ala Arg Ile His Arg Lys Gly Trp Ser Phe Cys Gln Lys  
 275 280 285  
 Leu His Glu Leu Val Leu Ser Phe Asn Asn Leu Thr Arg Leu Asp  
 290 295 300  
 Glu Glu Ser Leu Ala Glu Leu Ser Ser Leu Ser Val Leu Arg Leu  
 305 310 315  
 Ser His Asn Ser Ile Ser His Ile Ala Glu Gly Ala Phe Lys Gly  
 320 325 330  
 Leu Arg Ser Leu Arg Val Leu Asp Leu Asp His Asn Glu Ile Ser  
 335 340 345  
 Gly Thr Ile Glu Asp Thr Ser Gly Ala Phe Ser Gly Leu Asp Ser  
 350 355 360

Leu	Ser	Lys	Leu	Thr	Leu	Phe	Gly	Asn	Lys	Ile	Lys	Ser	Val	Ala
				365					370					375
Lys	Arg	Ala	Phe	Ser	Gly	Leu	Glu	Gly	Leu	Glu	His	Leu	Asn	Leu
				380					385					390
Gly	Gly	Asn	Ala	Ile	Arg	Ser	Val	Gln	Phe	Asp	Ala	Phe	Val	Lys
				395					400					405
Met	Lys	Asn	Leu	Lys	Glu	Leu	His	Ile	Ser	Ser	Asp	Ser	Phe	Leu
				410					415					420
Cys	Asp	Cys	Gln	Leu	Lys	Trp	Leu	Pro	Pro	Trp	Leu	Ile	Gly	Arg
				425					430					435
Met	Leu	Gln	Ala	Phe	Val	Thr	Ala	Thr	Cys	Ala	His	Pro	Glu	Ser
				440					445					450
Leu	Lys	Gly	Gln	Ser	Ile	Phe	Ser	Val	Pro	Pro	Glu	Ser	Phe	Val
				455					460					465
Cys	Asp	Asp	Phe	Leu	Lys	Pro	Gln	Ile	Ile	Thr	Gln	Pro	Glu	Thr
				470					475					480
Thr	Met	Ala	Met	Val	Gly	Lys	Asp	Ile	Arg	Phe	Thr	Cys	Ser	Ala
				485					490					495
Ala	Ser	Ser	Ser	Ser	Ser	Pro	Met	Thr	Phe	Ala	Trp	Lys	Lys	Asp
				500					505					510
Asn	Glu	Val	Leu	Thr	Asn	Ala	Asp	Met	Glu	Asn	Phe	Val	His	Val
				515					520					525
His	Ala	Gln	Asp	Gly	Glu	Val	Met	Glu	Tyr	Thr	Thr	Ile	Leu	His
				530					535					540
Leu	Arg	Gln	Val	Thr	Phe	Gly	His	Glu	Gly	Arg	Tyr	Gln	Cys	Val
				545					550					555
Ile	Thr	Asn	His	Phe	Gly	Ser	Thr	Tyr	Ser	His	Lys	Ala	Arg	Leu
				560					565					570
Thr	Val	Asn	Val	Leu	Pro	Ser	Phe	Thr	Lys	Thr	Pro	His	Asp	Ile
				575					580					585
Thr	Ile	Arg	Thr	Thr	Thr	Met	Ala	Arg	Leu	Glu	Cys	Ala	Ala	Thr
				590					595					600
Gly	His	Pro	Asn	Pro	Gln	Ile	Ala	Trp	Gln	Lys	Asp	Gly	Gly	Thr
				605					610					615
Asp	Phe	Pro	Ala	Ala	Arg	Glu	Arg	Arg	Met	His	Val	Met	Pro	Asp
				620					625					630
Asp	Asp	Val	Phe	Phe	Ile	Thr	Asp	Val	Lys	Ile	Asp	Asp	Ala	Gly
				635					640					645
Val	Tyr	Ser	Cys	Thr	Ala	Gln	Asn	Ser	Ala	Gly	Ser	Ile	Ser	Ala
				650					655					660
Asn	Ala	Thr	Leu	Thr	Val	Leu	Glu	Thr	Pro	Ser	Leu	Val	Val	Pro
				665					670					675
Leu	Glu	Asp	Arg	Val	Val	Ser	Val	Gly	Glu	Thr	Val	Ala	Leu	Gln
				680					685					690
Cys	Lys	Ala	Thr	Gly	Asn	Pro	Pro	Pro	Arg	Ile	Thr	Trp	Phe	Lys
				695					700					705
Gly	Asp	Arg	Pro	Leu	Ser	Leu	Thr	Glu	Arg	His	His	Leu	Thr	Pro
				710					715					720
Asp	Asn	Gln	Leu	Leu	Val	Val	Gln	Asn	Val	Val	Ala	Glu	Asp	Ala
				725					730					735
Gly	Arg	Tyr	Thr	Cys	Glu	Met	Ser	Asn	Thr	Leu	Gly	Thr	Glu	Arg
				740					745					750
Ala	His	Ser	Gln	Leu	Ser	Val	Leu	Pro	Ala	Ala	Gly	Cys	Arg	Lys
				755					760					765
Asp	Gly	Thr	Thr	Val	Gly	Ile	Phe	Thr	Ile	Ala	Val	Val	Ser	Ser
				770					775					780
Ile	Val	Leu	Thr	Ser	Leu	Val	Trp	Val	Cys	Ile	Ile	Tyr	Gln	Thr
				785					790					795
Arg	Lys	Lys	Ser	Glu	Glu	Tyr	Ser	Val	Thr	Asn	Thr	Asp	Glu	Thr
				800					805					810
Val	Val	Pro	Pro	Asp	Val	Pro	Ser	Tyr	Leu	Ser	Ser	Gln	Gly	Thr
				815					820					825
Leu	Ser	Asp	Arg	Gln	Glu	Thr	Val	Val	Arg	Thr	Glu	Gly	Gly	Pro

830	Gln Ala Asn Gly His Ile Glu Ser Asn Gly Val Cys Pro Arg Asp	840
845	Ala Ser His Phe Pro Glu Pro Asp Thr His Ser Val Ala Cys Arg	855
860	Gln Pro Lys Leu Cys Ala Gly Ser Ala Tyr His Lys Glu Pro Trp	870
875	Lys Ala Met Glu Lys Ala Glu Gly Thr Pro Gly Pro His Lys Met	885
890	Glu His Gly Gly Arg Val Val Cys Ser Asp Cys Asn Thr Glu Val	900
905	Asp Cys Tyr Ser Arg Gly Gln Ala Phe His Pro Gln Pro Val Ser	915
920	Arg Asp Ser Ala Gln Pro Ser Ala Pro Asn Gly Pro Glu Pro Gly	930
935	Gly Ser Asp Gln Glu His Ser Pro His His Gln Cys Ser Arg Thr	945
950	Ala Ala Gly Ser Cys Pro Glu Cys Gln Gly Ser Leu Tyr Pro Ser	960
965	Asn His Asp Arg Met Leu Thr Ala Val Lys Lys Pro Met Ala	975
980	Ser Leu Asp Gly Lys Gly Asp Ser Ser Trp Thr Leu Ala Arg Leu	990
995	Tyr His Pro Asp Ser Thr Glu Leu Gln Pro Ala Ser Ser Leu Thr	1005
1010	Ser Gly Ser Pro Glu Arg Ala Glu Ala Gln Tyr Leu Leu Val Ser	1020
1025	Asn Gly His Leu Pro Lys Ala Cys Asp Ala Ser Pro Glu Ser Thr	1035
1040	Pro Leu Thr Gly Gln Leu Pro Gly Lys Gln Arg Val Pro Leu Leu	1050
1055	Leu Ala Pro Lys Ser	1065
1070		

&lt;210&gt; 55

&lt;211&gt; 1315

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 095765CB1

&lt;400&gt; 55

```

gaagaagagc cgcgaccgag agaggccgcc gagcggtcccc gccctcagag agcagcctcc 60
cgagacagag cctcagcctg cctggaagat gccgagatcg tgctgcagcc gctcgggggc 120
cctgttgctg gccttgctgc ttcaggcctc catggaagtg cgtggctggt gcctggagag 180
cagccagtgt caggacctca ccacggaaag caacctgctg gactgcatcc gggcctgcaa 240
gcccgaacctc tcggccgaga ctcccatggt cccgggaaat ggcgacgagc agcctctgac 300
cgagaacccc cggaagtacg tcatgggcca cttccgctgg gaccgattcg gccgcccga 360
cagcagcgat ggtgccaaag cgggcccgcg cgagggcaag cgctcctact ccatggagca 420
cttccgctgg ggcaagccgg tgggcaagaa cggcgcccca gtgaagggtg accctaacgg 480
cgccgaggac gactcgcccg aggccttccc cctggagttc aagagggagc tgactggcca 540
gcgactccgg gagggagatg gcccgcagcg ccttgccgat gacggcgagc gggcccagcg 600
cgacctggag cacagcctgc tgggtggcggc cgagaagaag gacgagggcc cctacaggat 660
ggagcacttc cgctggggca gcccgcctaa ggacaagcgc tacggcggtt tcatgacctc 720
cgagaagagc cagacgcccc tgggtgacgtt gttcaaaaac gccatcatca agaacgccta 780
caagaagggc gagtggggc acagcgggct cccagggcta cctccccca ggaggtcgac 840
cccaaagccc cttgctctcc cctgccctgc tgccgcctcc cagcctgggg ggtcgtggca 900
gataatcagc ctcttaaagc tgctgtagt taggaaataa aacctttcaa atttcacctt 960
tccagaagtg gtgcacacga tacctgctcc gtcctcctca ctgaatttgt cctgagatca 1020
ggtgtggtcg tgaatattaa acatgcggat tgcaacccta gacagagctc ccttgagcgg 1080
ttgagcagat gcagccaggt gtggcgctccg gctgtggggc gaggggggtca cacggggccg 1140

```

agtggcttca	gcgacgagtc	catagggaca	tggttgaggt	cccgccgtgg	tgaggacaca	1200
gggggttgcg	gcaggtcagg	ccaatgcagg	gtccgcagtg	cggtgtaggg	tccactcatt	1260
ttgcgggggt	ggcgtctcat	tctcccat	gtctgccaag	ctgtaaacga	cggtta	1315

&lt;210&gt; 56

&lt;211&gt; 3796

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 6399886CB1

&lt;400&gt; 56

gctgctcccc	tcttcctaag	cggccccccc	tctccccggc	agcagaagaa	gggggtgggac	60
ccgggcccgg	tccgggaggg	ggccctggag	gaatggatgg	tggcgggaagg	gcggagcagg	120
ggcggggccc	gcggagactc	cacggggcgc	cccgggcgtg	aggcaccac	tctgggagca	180
cagagagctc	aggtagcctg	cctagatggc	ggcgcgcacc	ctgggcccgc	gcgtcgggag	240
gctgctgggc	agcctgcgag	ggctctcggg	gcagcccgcg	cgcccgccgt	gcggggtgag	300
cgcgccgcgc	agggcggcct	cgggaccctc	gggcagcgct	cccgcagttg	cagcagcagc	360
agcacagcca	ggctcgtatc	ccgcgctgag	tgcacaggca	gcccgggagc	cggccgcctt	420
ctgggggcct	ctggcgcggg	acactctcgt	gtgggacacc	ccctaccaca	ccgtctggga	480
ctgcgacttc	agcactggca	agatcggctg	gttctctggg	ggccagttaa	atgtctctgt	540
caactgcttg	gaccagcatg	ttcggaagtc	ccccgagagc	gttgctttga	tcctgggagc	600
cgatgagcct	ggaacggaag	tgaggatcac	ctacagggaa	ctactggaga	ccacgtgccg	660
cctggccaac	acgctgaaga	ggcatggagt	ccaccgtggg	gaccgtgttg	ccatctacat	720
gcccgtgtcc	ccattggctg	tggcagcaat	gctggcctgt	gccaggatcg	gagctgtcca	780
cacagtcatc	tttgctggct	tcagtgcgga	gtccttggtg	gggaggatca	atgatgcca	840
gtgcaagggt	gttatcacct	tcaaccaagg	actccggggg	gggcgcgtgg	tggagctgaa	900
gaaaatagtg	gatgaggctg	tgaagcactg	ccccaccgtg	cagcatgtcc	tgggtggctca	960
caggacagac	aacaaggctc	acatggggga	tctggacgtc	ccgctggagc	aggaaatggc	1020
caaggaggac	cctgtttgcg	ccccagagag	catgggcagt	gaggacatgc	tcttcatgct	1080
gtacacctca	gggagcaccg	gaatgcccac	gggcacgtc	catacccagg	caggctacct	1140
gctctatgcc	gccctgactc	acaagcttgt	gtttgaccac	cagccagggtg	acatcttttg	1200
ctgtgtggcc	gacatcggtt	ggattacagg	acacagctac	gtggtgtatg	ggcctctctg	1260
caatggtgcc	accagcgctc	tttttgagag	caccccagtt	tatcccaatg	ctggctcgga	1320
ctgggagaca	gtagagaggt	tgaagatcaa	tcagtctcat	ggcgcccca	cggtgtccg	1380
gctgttgctg	aaatacgggt	atgcctgggt	gaagaagtat	gatcgctcct	ccctgcggac	1440
cctggggctca	gtgggagagc	ccatcaactg	tgaggcctgg	gagtggcttc	acagggtggt	1500
gggggacagc	aggtgcacgc	tgggtggacac	ctggtggcag	acagaaacag	gtggcatctg	1560
catcgaccca	cggccctcgg	aagaaggggc	ggaaatcctc	cctgccatgg	cgatgaggcc	1620
cttctttggc	atcgtccccg	tctctatgga	tgagaagggc	agcgtcatgg	agggcagcaa	1680
cgtctccggg	gcctgtgca	tctcccaggc	ctggccgggc	atggccagga	ccatctatgg	1740
cgaccaccag	cgatttgtgg	acgcctactt	caaggccctac	ccaggctatt	acttacttgg	1800
agacggggct	taccgaactg	agggcgggta	ttaccagatc	acagggcgga	tggatgatgt	1860
catcaacatc	agtggccacc	ggctggggac	cgcagagatt	gaggacgcca	tcgccgacca	1920
ccctgcagta	ccagaaagtg	ctgtcatttg	ctacccccac	gacatcaaag	gagaagctgc	1980
ctttgccttc	attgtggtga	aagatagtg	gggtgactca	gatgtggtgg	tgcaggagct	2040
caagtccatg	gtggccacca	agatcgccaa	atatgctgtg	cctgatgaga	tccgtggtgt	2100
gaaacgtctt	ccaaaaacca	ggtctgggaa	ggtcatgcgg	cggtcctga	ggaagatcat	2160
cactagttag	gcccaggagc	tgggagacac	taccaccttg	gaggaccca	gcatcatcgc	2220
agagatcctg	agtgtctacc	agaagtgcaa	ggacaagcag	gctgctgcta	agtgagctgg	2280
caccttggtg	ggctcttggg	atgggcgggc	acccaagccc	tggcttgtcc	ttcccagaag	2340
gtacccctga	ggttggcgtc	ttcctacgtc	ccagaagcag	ccccacccc	acacatgacc	2400
cacaccgccc	tcacgtgaag	ctgggctgag	agccctttct	cccatccatt	ggaggtccca	2460
ggagtgtcac	ccatggagag	gctatgcgac	ctggctaggg	ctggttctgc	cagctgagtt	2520
tggtttccctg	gaatgaaaag	gcattgccat	ctccattcct	ctgccctctt	gagccagcac	2580
aggaagggtga	ggccctggga	tagcgcgcct	gctcagataa	cacagagcta	gttagctagt	2640
agcaaccgtg	ttttctccag	atctgtctag	atacaaaggt	cagaaatctt	atcttttatac	2700
ttttatattg	tggagaaca	gcattgcaaca	ctcacatgta	gtgtgtggat	ttacttgaac	2760
atgttctttt	taacatgtag	ttatgaaaat	ctcctttttt	gcctctactg	gtgaggaaac	2820
atgaggatca	gaggccacat	ttttaattat	tgttagtgta	tttgaagtc	tgaattggag	2880
atgtttgtac	ctctgtctaa	acagttccct	tgagaacttc	caagcctccg	gcattctttc	2940

ctggtgagtg tttctcctgt gcttggttgt gtataatgga gctaactcct aagcgggtggg 3000  
gtgaatgtgg ccgccttagt tctgaagcta ctccagttat gttctgttct ttcaagctgt 3060  
gatccagaaa gattttttgtg cccccagatg cctcttgata ggagaggcaa catactccaa 3120  
atagttgggt tcttcagga agctattaga aactcaggtg acttggttaga gcactaactt 3180  
ggtcagagcc aaatcctggc aaacgctgcc tgaccttcac tctgtggttg gggcagtgag 3240  
aaccactgag gtccaatgat gagacttga ggtctggatc cagtctctct ttgttttaatt 3300  
gtgacttagg tgctgtcaac attagcaaga taatggaaat cacgacgcca gtgggtgctt 3360  
acctccctgc taggcatgca ggggctggcg gttggcaggg gaaggaggcc cagtgaagccg 3420  
ggctcccttag gggagggaga gtttgcctc tttgccccac agtctaccct tcagggcctt 3480  
gtggcagtg cagtgttcgg ggggtgtctg ggccactgag taccactcg gtcgtggttg 3540  
tgctggcctc ttgggtgagt gaacctgtga agcccaggag gtggtgttgg ctgcagggtta 3600  
cacaaatact gagtgggtgt cttttgttac aggcttagca acaaagctgt gccctgggca 3660  
tggggggctg tagtgtagct acagttgtgc gtttgtgaaa tggcttagct ttccatgttg 3720  
ctgagaggaa cctggacatg gtcccgggca tctgaatgat ctgtagggga gggagttcaa 3780  
ataaagcttt attttg 3796

<210> 57

<211> 2983

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 6024420CB1

<400> 57

ccctgggagt ggccttggtt tcttgcagga cagccatgga cctactctgg atgccctctgc 60  
tgctgggtggc cgcttgtgtc tctgtgttcc acagctcacc agagggttaac gccggtgttt 120  
ccagcatcca cataaccaag cctgtgcaca tcttggagga acgcagctct ctagtgctaa 180  
cgcccgctgg cctgaaccag atgctgaacc agaccgctt cctcatggtg cttttccaca 240  
acctatctc aaagcaatcc aggaacttgg cggaagagct gggcaaagct gtggagatca 300  
tgggcaaagg caagaatggg atcggttttg gcaaagtgga cattaccata gagaaggagc 360  
ttcagcagga gtttgggatt accaaggccc cggagttag ctgttttttg agggcaacaa 420  
ggtcagagcc catcagctgc aaaggagtgg ttgaatctgc tgccttagtc gtttgggtga 480  
gacgacaaat tagccagaaa gcatttttgt tcaacagcag cgagcaggtg gcagagtttg 540  
tgatatccag gcccttgggt atcggttggct tcttccagga ttttagaggaa gaagtagcag 600  
agttgttcta tgatgtgatc aaagactttc cagagctaac gtttggagtc ataacgattg 660  
gcaatgtcat tggcgtttc cacgtcacc cctgacagct cctgggtgtt aaaaagggaa 720  
aaattgtgaa ccgcaaaaag cttattaatg acagtaccaa caaacaggaa ctcaatcgtg 780  
tcataaaaca gcaccttaca gattttgtga tcgaatacaa cactgagaat aaggatctga 840  
tttccgagtt gcacatcatg agtcacatgc tgctgtttgt ctccaaaagc tccgagtcac 900  
atggtatcat aattcagcat tataagctgg catcaaagga attccaaaac aagatccttt 960  
tcaccttgt ggatgcagac gaaccagaaa atggacgtgt cttcaagtac ttccgggtca 1020  
cagaggtcga tatcccattc gtccaaatcc taaacttag ctctgacgcc aggtacaaaa 1080  
tgccttcaga tgacataacc tacgaaagcc tcaagaatt tggccgcagc ttctagagta 1140  
aaaatgccac aaaacatcaa tccagtgaag agattccaaa atactgggac cagggtagg 1200  
ttaagcagct cgtggggaag aacttcaacg tagtcgtctt tgacaaagaa aaggacgtgt 1260  
ttgtgatgtt ctatgcaccc tgggtctaaa agtgcaagat gctgttccca ctgttgagg 1320  
aattgggcag aaaatatcaa aaccactcca caattatcat tgccaagatc gatgtcacag 1380  
caaatgacat tcagctgatg tacctggacc ggtaccatt cttcaggctg ttccccagcg 1440  
gctctcaaca agctgtcctg tataaggag aacacaccct gaagggtctt tctgacttcc 1500  
tggaagcca catcaaaact aagattgagg atgaggatga gctgttgtct gttgagcaaa 1560  
atgaagtgat agaagaggaa gtgctagctg aggaaaagga ggtgcctatg atgaagaaag 1620  
agttacctga acagcagtcg cctgagctgg agaacatgac caagtacgta tccaagctgg 1680  
aagagcccg cgtggaagaag aaaacatctg agggaggtgt ggtggtggtg gctaagccaa 1740  
agggacctcc agtgcaaaaag aagaaaccaa aagtcaagga agaactttag cttctccaat 1800  
accaggaaaa aagatgctta ttttccagat cctggcatca ttttctgaat ggattgattc 1860  
caataaaagc atatatcatt gtggtagggt aggtggggcg ggggtagggg tggataataa 1920  
agcctctgag tgtcaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1980  
aaaaaaaaaa aaaaaaaaaa aaacaacaca aacacacaaa caacaacaaa acacaaaaca 2040  
ccgcgggggg cgagcgaaga acaacaccac accgcacag aacaccacac cagagagaga 2100  
gagaggaatg gaacaagcac acaccaccaa caaacaaaaa aagagaagag aacacccgac 2160  
gagcgacaga agacggaaaa agaaacacac gacgcacgag cgaaacagaa agtcacaaca 2220

```

gaacaaaaaa gaacaaacac aaagcaacag agaaaacata agcagaaaga aaagcagaaa 2280
gaaaaggaga aacgagagaa acacacaaca gacaacggaa aaaaacaaaa gacaaaaaca 2340
ccaaatataa aaaaacataa aaataagaat aaaaagaaca aataaagaaa agaaaactaa 2400
aaaaaagcaa gagatagaaa agtaattgaa aaaaaaaga gtaaaatgaa caaacacagg 2460
aatagtacaa ctaaataata agacataata aaaccgcaa cagacacaac aaacaaagg 2520
aacacaacga gcaaacacac aacaaacaca acagccaaga acaacacaca gaagtacgag 2580
agagaaaagc gagagatagc gcagagacaa ccagaaaaaa gagaacagcc aaaaccaaga 2640
ggaaaacagg ataacaaaac ggaaagagag aagagggagg ccggagcgaa agagccgaga 2700
agacacgagc acagccgagg aacagcacga ccgaagaagc cgaagcgagg aaacagaacg 2760
agacaacaag gaaacagaag agaagtacag ccagagaaac gcagaacgac acatactagt 2820
gagaaagagg cgcatgtagc acaccacgag acaggagaag cacggtagca cgaacaagt 2880
cgagacgaag agcgaagaga gacaagcaga ggagaacaga cgagaaaaa aggaaaaagg 2940
agacaagaag gaacgacgag cgaggaagag cagagcggga gag 2983

```

&lt;210&gt; 58

&lt;211&gt; 3840

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7481067CB1

&lt;400&gt; 58

```

atgaaggcac ttttaccatt gacctttctg ttttttatta gttctccagg ttgggcaata 60
gataggcact gctacatagg cattgaagaa agcatttgga actatgctcc ttctggtaaa 120
aatatgctca atgaaaagcc tttttctgaa gacctagaat ttctacaagg aggtcaagcg 180
aggaagagct ttgtttttta aaaggctttg tattttcaat atactgataa tacatttcaa 240
aggatcattg aaaaaccatc ctgggtggga tttttagggt caatgattaa agcagagact 300
ggagacttca tttatgtaca tgtaaaaaat aatgcttcaa gagcttatag ttatcatcct 360
catgggctca cctactccaa agaaaatgaa ggtgctatct atcctgataa tacgacaggc 420
ctgcaaaaag aagatgaata tctggagcca gggaaacaat atacctacaa gtgggtatgta 480
gaagaacatc agggacctgg ccccaatgac agtaattgtg tgacaagaat ttaccattcc 540
catatagaca ctgcaagaga tgtagcttcg ggacttattg gaccaatact gacttgtaaa 600
agaggtacac tgaatggaga cactgaaaaa gatattgaca ggtcttcttt tctgatgttt 660
tctacaactg atgaaagcag aagctgggat agtgatgaaa atattcgtgc atttactgaa 720
tctggcaaga ttaatactag tgatccccgt tttgaggaga ctcaccatgt gtgctgaaga taggggtccag 840
aatggataca tctatggaaa tctgccaat ctcaccatgt gacatacacc ccgtctacct ccgcggaaca 900
tggtattttg ttggcatggg tggcgtggct gacatacacc accattatgc tcttcccctc ctcactggaa 960
actctgatct ctgggaatca cagaaggac accattatgc tcttcccctc ctcactggaa 960
gatgccttca tgggtggcca ggcccctgga gtgtggatgc tgggatgcca gatacatggt 1020
aagagtatgc aggcattttt caaagtaagt aattgccaga aaccttcaac agaagccttt 1080
gttactggga cacatgttat acattactat attgctgcta aagaaattct ttggaactat 1140
gctccatctg gtatagattt cttcactaaa aaaaatttaa cagcagctgg aagtaaatcc 1200
cagttatttt ttgaacgaag tccaaccaga attggaggaa ctaacaaaaa actgatttac 1260
cgtgaataca cagatgcttc cttccaaaca cagaaggcaa gagaagaaca ccttggaatc 1320
ctaggccccg ttattaaggc agaggtgaga cagaccatca aaatcacttt ctataacaat 1380
gcttccctgc cactcagcat tcagcctcct ggactgcatt acaacaagag cttagagggc 1440
ttattctacg aaacacctgg aggtaccctt ccaccctctt cacatgtaag tcttggcaca 1500
acatttgtct atacatggga agttccaaaa gatgtgggtc ccacctccac agatcccaac 1560
tgcttgacct gggtctatta ctcttcagta aatgggaaaa aagacatcaa cagtggcctt 1620
ctggggcctc tcttatatg tagaaatgga agtcttgagg acgatggcaa acagaaagga 1680
gtagacaaag agttttacct acttgccaca atatttgatg aaaatgaaag taatctcttg 1740
gatgaaaata tcagaacatt tatcacagag cctgaaaaa tagataaaga ggatacagac 1800
tgccaagcct caaataagat gtactccata aatggataca tgtatggaaa tctgcctgga 1860
ttggacacgt gcttaggaga caacgttttg tggcacgttt accttcaact ctttaggagc aagaagggac 1920
gattttacag ggatatattt ttcaggaaat accttcaact ctttaggagc aagaagggac 1980
acaataccta tgtttcctta tacttctcag acgtttttga tgacacctga ttctatagga 2040
acttttgatt tggtttgcac gacaataaag cacaatctag gaggcagtaa acataaatat 2100
cacgtgaggc aatgtgggaa gccaaacct gatcaaacac aataccagga ggagaaaata 2160
attattacca ttgcagccga ggaaatggaa tgggattatt ctcctagtag aaagtgggag 2220
aatgaactcc accacttacg aagagagcaa acgagcatgt atgtggacag aagtggaaaca 2280
cttcttgggt ccaaatacaa gaaagtctta tatcgtcaat atgatgataa cacgttcaca 2340

```



aatcaaacaa aaaggaatga aggtgaaaaa catctcgata tactagggtcc attaatattg 2400  
ctcaaccctg gtcaaataat tcaaattatc tttaaaaata aagccgcaag accgtattct 2460  
attcatgctc atggagtga aacaaataat tccactgttg ttccaactca gccaggagag 2520  
attcaaatat atacttggca gatacctgat agaactggtc ctacctcact ggactttgaa 2580  
tgcatacctt ggtttttacta ttcaactgta tctgtggcta aggaccttca cagtggactg 2640  
gtaggccctc tctctgtatg ccgcaaagac atcaacccca acatagttca ccgtgttctc 2700  
cacttcatga tatttgatga gaatgaatcc tggtaacttcg aagacagtat caacacctat 2760  
gcttcaaaac caaacaaggt ggacaaggaa aatgataatt ttcaactcag caaccaaatg 2820  
cacgcaatta acggaagact gtttggaat aaccaaggta taacattcca tgttggggat 2880  
gtagtgaatt ggtatctgat tggcataggg aatgaagctg acctgcacac agttcacttt 2940  
catggccata gctttgaata caagaatagg ggagtgtatc aatctgatgt ttatgacctt 3000  
cctcctgggg tctatcgaac tgtaaaaatg tatcgaagag atgttggaac ctggttattt 3060  
tattgccatg tttttgagca cattggtgct ggaatggaaa gcacttacac tgtacttgaa 3120  
agaaaagggc tgaaggagca gaacctctga agcagacaaa ggagagtcag catgaacagt 3180  
ttctcagaat cttctctcaa tatcaggact acatttgtca acaaaaccaa aaactgatta 3240  
gccaccgata taattttttac ctacaacatc ctattaatgt caataatata attattgata 3300  
caattctaata aatcactacc cttattccta tcagtgttca tgtacattct tagtaaaaga 3360  
gactttgggtg cgctgtccat gaaataaatc cccattgct aacattcttt ctttggaaaa 3420  
gtagattttg catttcaaag aatataaagt caaattggat tggatttaca ggtcatctgt 3480  
tcccacagaa ggggtgatatt gatgttgcta ttgataagta aactttttgt ggcaaaagtg 3540  
atggtagtta ttttaaggat gttocaaaga ctaataataa ttttgtattt attccttaaa 3600  
tgtatgtaat catttttagct tagtatttta acttagaact gcatgctatt atataatatt 3660  
acctattttt gaaacttcct tttctacagc ataaatattt gatatgatat gaatatggac 3720  
aagcttacaa gccaaagtaa agctgccaaa gaaggaaaac tccagggacc aaggagtctg 3780  
ggaggaacca gctaaagact ttcattgacaa tgtaccaggg agactagttt gagatcaagg 3840

&lt;210&gt; 59

&lt;211&gt; 1570

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3378720CB1

&lt;400&gt; 59

gagaacgaag ctggttggaa cgttggaaag tgctctctga ctacacttca caagcaaggg 60  
gcaccttttg tggactgaca ttccagaaag ggatgttgtg aaacaaaagc tgacatttat 120  
atatatatac atatatagc tatttgagtt cctcagtaga aagctatcat atatactcag 180  
aatgttttgg acgttttaag aatggttctg gttggaaaga ttctggcttc ctccaacaat 240  
aaagtggtea gatcttgagg atcacgatgg actcgtcttt gtaaaacctt ctcatattata 300  
cgtgacaatt ccatatgctt ttctcttgct gattatcagg cgtgtatttg aaaaatttgt 360  
tgcttcacct ctagcaaaat catttggcat taaagagaca gttcgaaagg ttacacacaaa 420  
tactgtctta gagaattttt tcaaacattc cacaaggcaa ccattgcaaa ctgatattta 480  
tggactggca aagaagtgtg acttgacgga gcgccagggtg gaaagatggg ttaggagtcg 540  
gcggaatcaa gagaggcctt ccaggctgaa gaaattccag gaagcttgct ggagatttgc 600  
attttactta atgatcactg ttgctggaat tgcgtttctt tatgataaac cttggctata 660  
tgacttatgg gaggtttgga atggctatcc caaacagccc ctgctgccat ccagtagctg 720  
gtactacatt ttagaaatga gtttttattg gtctctgtta ttttagacttg gctttgatgt 780  
caagagaaag gattttctag ctcatatcat ccaccacctg gctgctatta gtctgatgag 840  
cttctcttgg tgtgctaatt atattcgag tgggaccttc gtgatgattg tacacgatgt 900  
ggctgacatt tggctggagt ctgctaagat gtttttctt gctggatgga cgcagacctg 960  
taacaccctg tttttcatct tctccaccat atttttcatc agccgcctca ttgtttttcc 1020  
tttctggatt ttatattgca cgctgatctt gcctatgtat cacctcgagc ctttttttcc 1080  
atacatcttc ctcaacctac agctcatgat cttgcaggtc cttcaccttt actggggtta 1140  
ttacatcttg aagatgctca acagatgtat attcatgaag agcatccagg atgtgaggag 1200  
tgatgacgag gattatgaag aggaagagga agaggaagaa gaagaggcta ccaatggcca 1260  
agagatggat tgttttaaga acggcctcgg ggttgagagg cacctcatc ccaatggcca 1320  
gcatggccat tagctggaag cctacaggac tcccatggca cagcatgctg caagtactgt 1380  
tggcagcctg gcttccaggc cccacacaga cccacattc tgcccttccc tctttctcac 1440  
caccgccttc cctccacct aagatgtgtt taccaaaatg ttgttaactt gtgttaaat 1500  
gttaaatata agcatgcca tggattttta ctgcagttag gactcagact ggtcaaagat 1560  
ttcaaagatt 1570

<210> 60  
<211> 409  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 938824CB1

<400> 60  
cagtcttttc cccactttac cagtgtgtga ggctctctc ccacttctc cacttaccac 60  
ctctccaaga tgccagcctc actgtgggct ttcctagaa agaaacactg gtttctttct 120  
atcgtgccct ggtagtggt gtttctcaca ttaggcctct gtgttagaaa taaagctgct 180  
aaactccatg tcgttatata acaaaaggaa tacagtgacc tctcctcat tcttctgata 240  
gttccctcaa ctccagctgc agcccctgcc aaatactatc atccttaaaa gatagacagt 300  
gatcccagca ctgtgggagg ccaaggcagc tagatcactt gagtccggga gttcaagacc 360  
agcctgggca acatggtgaa ccccatctct actgagaaaa ttataacaa 409

<210> 61  
<211> 953  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1683721CB1

<400> 61  
aagagcgatt ataattgagt atgatgtgtg cagtaagagc aacacaagat gatgatagtg 60  
gtgggtggtac ttgtcatttg cttgatgcc a ggcacacttc taggtgctta catgcacctt 120  
ctcattgaat tccccaacag tctatgaag taggctcttt tcatcccat t gatagatga 180  
agaactccag gcccaaactg gtttaagtta ttgggttaaaa gtcacacaga aaatggccaa 240  
gccaggattt gaacctataa tctctgctct gcccttaaga gatagtacaa actggcggt 300  
tcgggaagcc tcacagagaa ggcagcattt gaaccaggac tgaaggatca ataagatgaa 360  
ttctggcatc aggaaaggaa aggcactccc tgcagtagga atgggagggg caaaggcaga 420  
gaggcgggac catggctcag catggtcact gattaggtga gtgtggctgg aaccgtgaca 480  
gggaacagca ggagatgatg ctgggttogg gatggaaggc tctgtcctg aagagccttg 540  
ctttccccac tcaggggtat cctgagggct atgaggagct acttaggaaa gtgacaggag 600  
cagatttgac ttggtcacct ggagatgcaa tccaattcca ggttctctggc accaggaaga 660  
caaagcagta ttgtgaattt gaaaatgaaa tcaactttat catgccccac atgaaaattc 720  
agtcgctctt atttttgctt ggcttttatg taaaagacct aagccaatga aactgcctgc 780  
cattagtcaa ggtcagagtg aaacttgtca gaaaaacttc cttaggtctc tctactgacac 840  
tacaagttat tgccaacctg agagctcctc cacagaaatt accatttgga gactgtccac 900  
agtgggattt cagataggct ccaacccctc acgggaacct ccccccccg cca 953

<210> 62  
<211> 890  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1694122CB1

<400> 62  
accacgcgtc cgcgagcgcg tgggtggcca aagtgcagga ttaccggcag gccaggtcg 60  
agaggctgga gaccaagggt gtcaaccccc tgaagctcta cggggcacag atcaagcaga 120  
cacgggctga gatcaagaaa ttcaaacatg tccaaaatca tgagatcaaa caactggaaa 180  
aactggagaa actgaggcag aagtcaccct cggatcagca aatgatctcc caggcagaga 240  
ccagagtgca gagggccgct gtggactcca gccgcaccac cctccagctg gaggagactg 300  
tggatggctt ccagaggcag aagctcaagg acctgcagaa atttttttgt gactttgtaa 360  
ctattgagat ggttttccat gccaaagcgg tggaggtgta ttctagcgcc ttccagacct 420  
tggagaagta tgacctggag agggatctac tggattttag agccaagatg caaggagttt 480

atgggcatta tgacactcgg ctgcttgcca acaccagccc cctccatct gttcttcagt 540  
ctctcgccag ccagagtgtc cagagcacca tatggagccc aggaaaagaa ggggaggaga 600  
gtgaggacaa ctccatggag gaggccccc tggaggacct cagggcactg gggcaggac 660  
cccataagag agaactgccc acaacagtca gaagaactta gctggccttg gatcctcagg 720  
tggtgtctgc tgtgtgccct caggcaagcc acgtgtcctc tgagcctcag tttcctcctc 780  
tgtacaacag ggccaatatc actcacttca caggttgctc tgggggatcg ctgtgcctgg 840  
catatagtag gtgttcaata aatgcctgt gactctcaaa aaaaaaaaaa 890

<210> 63

<211> 1960

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 1970615CB1

<400> 63

cggtcgagca caggcagggtg gtcaaaacaa ctttcaacca gaatctactg atatggatag 60  
atgctgtccc ttatggaggg acgcagactt cagaactgtg cccatgactg ctgactgcca 120  
ccaccaaggc cctcagggtac acagcctgcc tccctgagga tgatgtgggc agatttcagc 180  
cctagttaac agaagagtcc cccaggagggt aggggggcccc catcactgga gacatgccag 240  
cagagcctct ggccaccta ccaggagggtc ttctgaaatg actatacgag gtaaagaagt 300  
agtaccagat ggtcccaaag ttccctttta gctgaaagc ttttctttgt cctccttag 360  
tgaatctgtg ttccgagccc tactctaaag ttcaagtgtc aatacaatag tccaccaaga 420  
gactgggaat gattagaagt gaaattgggt cctccttacc aaggaggggc agatgatctc 480  
cattgcacag ggcgattaga ttctggagct gaggtgggga ctgcaggagg ccacctagtc 540  
tggtagggtt caaccaagc tgtgtacatt agaattccct tgggagcgtg caggaaatac 600  
agatgcccac gccacattcc agaccaactg aagctgaatc tccagagtag ggcctgtatg 660  
ttcatataag ctccacaggt gatctgcagt acagtgaaga tggagactg catgtgtacc 720  
tatttgcaat aaagatgaag aggacagcaa gctccagaca ggagctggga ctcaaccag 780  
atctcttaag tctgcctgg tggctcctta aaagtccaga agtgttgccc caagccctcc 840  
ctcaacatct ctgggaaccg cagctgcagc acgatggggg ttcagtgcc ctgtttgccc 900  
cttaccagc tgtggtttat tctgcttgta tgtctgcaca ggccggatgc tctgttccct 960  
tgtcttattc tccatttact cagtactgg ggctcactcc cgtctgatgc actagccaag 1020  
attgccttag tgtgctccag aaaagaaggc caaatcccag gcattgtcag ggcagcagag 1080  
ctctacagga taggcttacc tttcccacct gtgtggctag cacttcacag tttacaaatt 1140  
cctccacct ccaactcagt acacatgctg ttctaacaca ggtcaggcag gcattacagt 1200  
ccccatgttc agaatacaag acctagcctc agagaagtga agaaacatca tgccaagggtc 1260  
attgactgcc aagcggtaga ggtggggtg catccagaga gcttcccggg atgacctctgc 1320  
acaatgccat tcttggtcca gctccctcca cccaaggga cccagactgc acacttaaca 1380  
aacaggacac aggtgtcttt gaacaaactt ttttgtatta ttatttttac atctagaata 1440  
aattatttaa attatttcac agcaaggag agggataggt aatttttatc agatattttt 1500  
ttaaacatc tgttttttaa attacatttt tgtttatgtt cttgagctga tgtagtggaa 1560  
cttgccctagc acattcaggt cccagccagt tggcagagca tgctctcctc tcttattcc 1620  
ataccctggg cgtcccttt ctgttgactc aggaactttc tgagaatgag gacagcacta 1680  
ggagatgagc tttggcagggt atccacctta acgctacaat aattgtgctt cctgaaacaa 1740  
aacttgagat tgtatcatag aaggaaacag gaagtcagaa atcaaatact tgcttttaatt 1800  
tgaaaccgtg cctgaaacag tttgaatgat tgttttaatt ttgtttctga aattccttgt 1860  
acctttgtga aaaataatga taataaataa aagtgaataa aaatagatgt ggaatatgca 1920  
atggaaataa tgtaacaaaa taataaacat ctgcaagtag 1960

<210> 64

<211> 832

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2314152CB1

<400> 64

ggtttttttt tttttttttt ttttttttaa atgtcagtat ctatatttaa tttcattgct 60

```

tcaaaacata tatatgtatg tatatacata aattttaaaaa ctgacagaat gacaaagatt 120
tttcagcaat acataatcat ggtgggagag ttttaacatac ctctctcaat aaatgctaga 180
tcacagatca acatttataa atgtgtagat ggttttttaa caccatttgct tgatctagca 240
gacatttgta gaatacaata cattgatttc aaccatccat gaagcattta taaaaactga 300
catatactaa gtgcatcagt taggaattat attcaggtga aataaacctc ccaccaaaaa 360
caaaaacaaa caaaagaaaa ctgtaccagt aagaagtga tatttttcat gtatgggaca 420
tccagaaata aatagtcag ggctgctgca atcacataaa tatgccatt ctattaaatc 480
tattcctaaat attttttaga ttacagaagc aagatgggta tgacttccgg ccaccctctt 540
cttagcttgc ggcttttacc cttatgggta caggagggct cctctagggtc tagaaatcat 600
gtctacttgt ccaaaaggca agaagtggag agatgtggct acatgaaacc atccctgaac 660
acaatatctt ccccagagtc acatccagtg acttctcata ttcatactag ccaggaccgg 720
aggaaatggc ctgcccttgc ttgcaagaag ggctgggaga tggaagcttt tttttattat 780
tattattttt gagacagagt ctcatgctgt caccagggt ggattgcagt gg 832

```

&lt;210&gt; 65

&lt;211&gt; 546

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2886225CB1

&lt;400&gt; 65

```

caatgtacga gctcggcatc atagtacggg ccgcagtggtg ctggaaagga aatcacagcc 60
tgagtgtgtc tcagtcagcg cagcagaatt cagcgcgga aagctgaatg caaccctgg 120
tgcaggaagg gaggcttttc ctgaggaccg ggagaggatt ttaagtacat agaaggaagc 180
ttctggagat cctgctccgt cgccccagtg ttcagactac ctgttcagga caatgccgtt 240
gtacagtagt ctgcacattg gttagactgg gcaagggaga gcaacgccat ggaccgctgg 300
ggacaaaatg ggctggttcc aaggagaaga catttggttg ctctttttt gaattcctata 360
tccagtgttt ttcttcacag gttttgcaca ctagggacca agaagccctc ggggacactg 420
ctgagaaaaa actgccgaag ggaagaccag cgggagatat acaaatattt tagggatcat 480
ggaatctatt ccagaggaaa ctaaccaaca gtgagcatct tcaactaaaa ctaccctgg 540
atgtaa 546

```

&lt;210&gt; 66

&lt;211&gt; 890

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 6144418CB1

&lt;400&gt; 66

```

gtacttttga atcaagcagg tgttacagat ctctctctcc tagggagctt gcagtatttt 60
accagtttg tggagttcca tttgaggatt gctgggggtt gtggaccatc cagagatctt 120
tccatggtaa agatcaagca tctggtgttg ccatgatcta ttgtcagatt gagtctcact 180
ctgttgccca ggctggagca cagtggcgtg atcttggtc actgcaacct ccagggttca 240
agggaaacctg taaaggaaac cagttagctta tctgtaccag tggcatatct gtataatctc 300
ttagaaaatc aagggttcaa aatcaaacac cttgctagtt tgcgattcag tgattcaaga 360
atgaattctg cagttggagg actctctagg cccttcagtg ttccacttac tttttctgct 420
cttattccct ctcttcttct acacgccagc gtctcttct gcactggatg gtatcatgat 480
tttcaggaag gagagtcaaa aagggaaca agtcagctca agcaaaaaca tcttggcaca 540
cgggaggagc aagtaataaa tgatagtagt tgggacacca ttagtcattg tcattctgca 600
tgctccagca ccaataaaac catcttaacc aaacaccctt ggataattgg ttcccatgac 660
taatacatag ccactgatcc agagcctaag gagtgtgaat gacaaacaat aaacctgctt 720
gggataacgt gattgattgt aattccattt gaaataaatg taacaaacag caatgagcca 780
tctagttaac catgcagcac aacttaattg ttacggctgt tcatttcatt cccagtgctc 840
ctttccagaa acacaaaaaa aggggcggca aaaggaggcg gagaccggat 890

```

&lt;210&gt; 67

&lt;211&gt; 807

<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 6834184CB1

<400> 67  
agagagagaa acataagggga attagggagg gtctctgggg tggtgacatt taagccaaag 60  
ccagaaggac aggaaggaca cagctacgca gacagctgga ggaagagcat tccagacaga 120  
gggaactgtg tactcacagg ctccaaggaa gaacagaact aggtgttgga gggattgagg 180  
ggaaccacc tctgtgtgtg gcttcgggtc ccggacaagg gggctgagat agaacgaaga 240  
gggagaggag ggcagagcct gggatcatga gggagttccc tgcctgggct ggctgctctc 300  
tagtgctttt tccctcatgt cctgggggag tctgcacggg tgtgccctgt tgttgccatt 360  
gtgctcaggg acctttgaag ttgagaaaat actagtgggt gtgggggctg acgagtgcc 420  
ggcatcagcc ctggtctggg aggtaccat gctcacctc cagctgcacc cacggggctc 480  
cacctcgag cctccagagc cagactgctc tgctgcagt ctgggcaaat tgtaacctt 540  
tctgtgcctt agtttcttca tctgtgaatt gggggttaata gcatccaatg agagcaaagg 600  
gcttggtaca gtaactaagc tttggttagt atgagcaaag ggctgagtgt agacatagt 660  
cactcactcc gtgagcattg ttccctgtggc tgtggatggc tccgtggatg tgtgtgtttc 720  
tagcatagag gacagaggct gtctccagac ccggaccctg acttgccctg tgctcctgtt 780  
ggcactcatt agcctcttgt gctcacc 807

<210> 68  
<211> 677  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 6951005CB1

<400> 68  
attgtcacat gtttttttgt ttgtttgttt gtttgtttct gtttttgttt tgctgtggat 60  
gagttaaaaa gttgggaagt ttctgtttta gtttgtcttt ctgtactgcc aagctgggaa 120  
gtaaaaaatgc cacttcccat tcctctataa ctaacacaac atttccaaaa tcatagtaag 180  
actatcttct gaaaactgag ctctctccca aattctctac attgctaatt gttttgcccag 240  
ggttcccatc agtcccctct cccctgcccc atcctctgtg gcttcttccct ctggccccat 300  
ccatactgga tcaattctca ctagggtccca ctctgagatc tccagcattc attccctctc 360  
gtgattctcc tgcctcaatt gcagttacag atattactat tcatatccag atagtacttc 420  
tagtactct tctggcctct agtttcacaa agtccccga cttcagttac aatcctgatc 480  
tgtcatttac cagcagctat atgacctcag gaatgttggt ggacatttct gagctgcaat 540  
atccctatgt gcaaagtga actatttaaat atttttctca cagggctgtt ctccagaatca 600  
gaggaatatg gagataaata tatatatata catataaagc tggtaaggta ccaagggcta 660  
gcacagtaga gccagaa 677

<210> 69  
<211> 617  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 7250331CB1

<400> 69  
tgggctgcac cactcacaga gctccctccc ccaggcactt agttggggcc cagcactgac 60  
ctttcccttg agcccaggat gtggccagag cccctctggt gacctctctc gcccttctc 120  
tgctctctca gcttgagctg cctgcccga gttcggctgt tccggggcca gtgtgtcacc 180  
tgccaacttc cacatcacc tcctccctcg ctccctctc tccttcccca aggacctccc 240  
cccatttctg gcagccaagc cattaatctg gagacagaaa tgggtttgct atcgattctc 300  
tgccacttt ttctttcatt acaatttgta ccgtgattct tctcaccctt ctctgcgtcc 360  
atgcatttaa agagttgtct ctttaaattg tgaagcttcc ggaagcctga tgctattctg 420

```

tgtctccttt caaaggaaga agggggggcc cagctatggg tgaggactca agttattagt 480
ttggaaatag agcaactatg tgtacagccc accttagagg tcatgttacc cccttcctgt 540
taaattttac aattaatttt ggttcaggaa atgtaaataa atttgtaaat tacaaatagc 600
aaaaaaaaaa aaaaagg                                     617

```

<210> 70  
 <211> 795  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1758413CB1

```

<400> 70
atggctccag gctcccggac gtccctgctc ctggcttttg ccctgctctg cctgcccctgg 60
cttcaagagg ctgggtgccg ccaaaccgtt ccgttatcca ggctttttga ccacgctatg 120
ctccaagccc atcgcgcgca ccagctggcc attgacacct accaggagtt tgaagaaacc 180
tatatcccaa aggaccagaa gtattcattc ctgcatgact ccagacctc cttctgcttc 240
tcagactcta ttccgacacc ctccaacatg gaggaacgc aacagaaatc caatctagag 300
ctgctccgca tctccctgct gctcatcgag tctgtggctgg agcccgtgcg gttcctcagg 360
agtatgttcg ccaacaacct ggtgtatgac acctcggaca gcgatgacta tcacctcta 420
aaggacctag aggaaggcat ccaaaccgtg atgggggtga ggggtggcgcc aggggtcacc 480
aatcctggaa cccactggc ttcgagggct gggggagaga aatactgctg ccctcttttt 540
agcagtaagg cgctgaccca agagaactca ccttattctt catttcgctt ggtgaatcct 600
ccaggccttt ctctacaccc tgaagcggag ggaggaaaat ggataaatga gagagggagg 660
gaacagtgcc caagcgcttg gcctctcctt ctcttccttc actttgcaga ggctggaaga 720
cggcagccgc cggactgggc agatcctcaa gcagacctac agcaagtttg acacaaactc 780
gcacaacctat gacgc                                     795

```

<210> 71  
 <211> 1677  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7011042CB1

```

<400> 71
ggggagggaag agcaggggggt atcacctgct tcttagggga gacgacaggg caaaagcaga 60
caggggagag gtcactgcac actaggcgac tctgggaacg tctccccgcc ctgcagggaa 120
catccagcgc ctgtgctcct cctcagagcc ctggaggtag ggttgggacg cgcactctgac 180
tctttgtgcg ggggttccctg tggatttgat gggcgctcctg ctgttctctc ccagacgcc 240
atgcgggcaa ccctaccgct gctgctgctg acggtgctgc gcccagctg ggcagacct 300
ccccaggaga aggtcccgc tttccgggtc actcagcagg gcccctgggg gagcagtggc 360
agcaacgcca cggactcgcc ctgcgagggg ctgcccgcgc cggatgcgac ggccttgacc 420
ctggcgaaac gcaacctgga gcgcctgccc ggctgcctac cgcgcacact gcgcagcctc 480
gacgccagcc acaacctgct gcgcgccttg agcacttccg agctcggcca cctggagcag 540
ctgcaggtgc tgacctgctg ccacaaccgc atcgccgcgc tgcgctgggg cccgggtggg 600
ccggcggggc tgcacacctt ggaacctcagc tacaaccagc tggccgctct gccgccgtgc 660
accgggcccc cgctgagcag cctccgcgcc ctggcgctcg cggggaatcc gctgcggggc 720
ctgcagcccc gggccttcgc ctgcttcccc gcgctgcagc tctcaacct ctctgcacc 780
gcgctgggtc gcggagccca ggggggcata gccgagggcg cgttcgctgg agaggatggc 840
gcgcccctgg tcacgctcga agtcttgat ctacagcgga cgttccttga acgggttag 900
tcagggtgga tcagagacct gccgaagctc acatccctct acctgaggaa gatgcctcgg 960
ctgacgacct tggaggggga cattttcaag atgaccccca acctgcagca gctggactgt 1020
caggactccc cagcacttgc ttctgtcgcc acacacatct ttcaagatac tccacatcta 1080
caggtccttc tgttcagaa gtaagtgcct ctgaggcaca tcttcatcac atgactgatt 1140
tttgccatt acgctatggt gaattttata gaacaaacca gaccttaatt tttctccac 1200
tactcttcca aatcctctct ggggctttgt tttccccac cctttgggtg atatcttggg 1260
taggtcccag atagtccacc cagcccatct agccttgcaa ctcagtcagc ttcagagtag 1320
cattagaacg caaaactcct atgcctttaa agtttatttt aagcccagtc ttagaaatct 1380

```

acagtgggaa gaggtgagga ggacctcagg tctcctctct gtctggatct gctttcttcc 1440  
ccacatgtca tgcaagttca gtcccttcca gaatttgagt ggggtctaggg atgaaagtat 1500  
tgatattgtt agaaaatccc ttggaagtct atgggcaggc cctgttagtc ctcttttaca 1560  
caatgcagtc actaaagatg gaattattct tctaaaagga tgaacaactt ttcaaagcaa 1620  
aggcagaacc attcactcat tctgtccttt attcaataaa tagctattga gcacaaa 1677

<210> 72

<211> 1402

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7427362CB1

<400> 72

ggcgcgcgcg gctcggggttt cggcccgcgc cggcgcgccg cgtgatcccc tcccgggcgc 60  
ggggcggggc cgggcccagc tggccgcgct cggggcgcta taagagggcg gcggcgcgcg 120  
cgcgccctgc gcggagctgg gagcgcgatg gtcggcgccc gaggcgcggc aagatgctgg 180  
atgggtcccc gctggcgcg cgggtggccg cggccttcgg gctgacgctg ctgctcgccg 240  
cgctgcgccc ttcggcgccc tacttcgggc tgacgggcag cgagcccctg accatcctcc 300  
cgctgaccct ggagccagag gcggctgccc aggcgacta caaggcctgc gaccggctga 360  
agctggagcg gaagcagcgg cgcattgtgc gccgggaccc gggcgtggca gagacgctgg 420  
tgaggggcgt gagcatgagt gcgctcgagt gccagttcca gttccgcttt gagcgtctga 480  
actgcacgct ggagggccgc taccgggcca gcctgctcaa gcgaggcttc aaggagactg 540  
ccttctctta tgccatctcc tcggctggcc tgacgcacgc actggccaag gcgtgcagcg 600  
cgggcgcgat ggagcgctgt acctgcgatg aggcaccgca cctggagaac cgtgaggcct 660  
ggcagtgggg gggctgtagc gaggacatcg agtttggtgg gatggtgtct cgggagttcg 720  
ccgacgcccg ggagaaccgg ccagatgccc gctcagccat gaaccgccac aacaacgagg 780  
ctgggcgcca ggtgatcaag gctgggggtg agaccacctg caagtgccac ggcgtgtcag 840  
gctcatgcac ggtgcggacc tgctggcggc agttggcgcc tttccatgag gtgggcaagc 900  
atctgaagca caagtatgag acggcactca aggtgggcag caccaccaat gaagctgccg 960  
gcgagggcag tgccatctcc ccaccacggg gccgtgcctc gggggcaggt ggcagcgacc 1020  
cgctgccccg cactccagag ctggtgcacc tggatgactc gcctagcttc tgcctggctg 1080  
gccgcttctc cccgggcacc gctggccgta ggtgccaccg tgagaagaac tgcgagagca 1140  
tctgctgtgg ccgcggccat aacacacaga gccgggtggt gacaaggccc tgccagtgcc 1200  
aggtgcgttg gtgctgctat gtggagtgc ggcagtgcac gcagcgtgag gaggtctaca 1260  
cctgcaaggg ctgagttccc aggcctgcc agccctgctg cacagggtgc aggcattgca 1320  
cacggtgtga aggtctaca cctgcacagg ctgagctcct gggctcgacc agcccagctg 1380  
cgtgggggtac aggcattgca ca 1402

<210> 73

<211> 1251

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7485304CB1

<400> 73

atgcttgctg tggatgatggc tgatttggtt tccctgatgt gctgggtctg caagcagaaa 60  
ctgccaggct tggcagcctg gtctgcggct gtgagacagg aagtggggct gtgcttgag 120  
agacaaagcc tacagctgga cccggctctt tctctctga gtcagggatg gcccctgagg 180  
aggccccttc ccttcatttg cccctcacca ccatecccaa ggctcacctg tctccctcct 240  
ctcgctctct ctacgctgac cgggcgggaa gtcctgacgc ccttcccagg attgggcact 300  
gcggcagccc cggcacaggg cggggcccac ctgaagcagt gtgacctgct gaagctgtcc 360  
cggcggcaga agcagctctg ccggagggag cccggcctgg ctgagaccct gagggatgct 420  
gcgcacctgc gctgcttga gtgccagttt cagttccggc atgagcgctg gaactgtagc 480  
ctggagggcg ggatgggcct gctcaagaga ggcctcaaag agacagcttt cctgtacgcg 540  
gtgtcctctg ccgcctcac ccacaccctg gcccgggcct gcagcgtggc gcgcatggag 600  
cgctgcacct gtgatgactc tccggggctg gagagccggc aggcctggca gtggggcgctg 660  
tgcggtgaca acctcaagta cagcaccaag tttctgagca acttctggg gtccaagaga 720

```

ggaacaagg acctgcgggc acgggcagac gcccacaata cccacgtggg catcaaggct 780
gtgaagagt gcctcaggac cacgtgtaag tgccatggcg tatcaggctc ctgtgccgtg 840
cgcacctgct ggaagcagct ctccccgttc cgtgagacgg gccagggtgct gaaactgcgc 900
tatgactcgg ctgtcaagggt gtccagtgcc accaatgagg ccttggggccg cctagagctg 960
tgggccccctg ccaggcagggt cagcctcacc aaaggcctgg cccaagggtc tggggacctg 1020
gtgtacatgg aggactcacc cagcttctgc cggcccagca agtactcacc tggcacagca 1080
ggtagggtgt gctcccggga ggccagctgc agcagcctgt gctgcgggcg gggctatgac 1140
acccagagcc gcctgggtggc cttctcctgc cactgccagg tgcagtgggtg ctgctacgtg 1200
gagtgccagc aatgtgtgca ggaggagctt gtgtacacct gcaagcacta g 1251

```

&lt;210&gt; 74

&lt;211&gt; 4961

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1422394CB1

&lt;220&gt;

&lt;221&gt; unsure

&lt;222&gt; 4929

&lt;223&gt; a, t, c, g, or other

&lt;400&gt; 74

```

ccgtcttcat cttgcgaaca cttgcgagac cgtcgctaata gaatcttggg gccgggtgtcg 60
ggccggggcg gcttgatcgg caactaggaa accccaggcg cagaggccag gagcgagggc 120
agcgaggatc agaggccagg ctttcccggc tgccggcgct cctcggaggt cagggcagat 180
gaggaacatg actctcccc ttcggaggag gaagggaagtc ccgctgccac cttatctctg 240
ctcctctgcc tctctcctgt tcccagagct ttttctctag agaagatttt gaaggcgggt 300
tttggattct tcacttctct tgaacaagga actcactcag agactaacac aaagggaagta 360
atttcttacc tggctattat ttagtctaca ataagttcat ccttcttcag tgtgaccagt 420
aaattcttcc catactcttg aagagagcat aattggaatg gagagggtgct gacggccacc 480
caccatcacc taaagaagat aaacttggca aatgacatgc aggttcttca aggcagaata 540
attgcagaaa atcttcaaag gaccctatct gcagatgttc tgaatacctc tgagaataga 600
gattgattat tcaaccagga tacctaattc aagaactcca gaaatcagga gacggagaca 660
ttttgtcagt tttgcaacat tggaccaaata acaatgaagt attcttgctg tgctctgggt 720
ttggctgtcc tgggcacaga attgctggga agcctctgtt cgactgtcag atccccgagg 780
ttcagaggac ggatacacga ggaacgaaaa aacatccgac ccaacattat tcttgtgctt 840
accgatgatc aagatgtgga gctgggggtcc ctgcaagtca tgaacaaaac gagaaagatt 900
atggaacatg gggggggccac cttcatcaat gcctttgtga ctacacccat gtgctgcccc 960
tcacgggtcct ccatgtctac cggaaggtat gtgcacaatc acaatgtcta caccaacaac 1020
gagaactgct cttccccctc gtggcagggc atgcatgagc ctcggaactt tgctgtatat 1080
cttaacaaca ctgggtacag aacagccttt tttggaaaat acctcaatga atataatggc 1140
agctacatcc cccctgggtg gcgagaatgg cttggattaa tcaagaattc tcgcttctat 1200
aattacactg tttgtcgcaa tggcatcaaa gaaaagcatg gatttgatta tgcaaaggac 1260
tacttcacag acttaatcac taacgagagc attaattact tcaaaatgtc taagagaatg 1320
tatccccata ggcccgttat gatgggtgatc agccacgctg cgccccacgg ccccgaggac 1380
tcagccccac agttttctaa actgtacccc aatgcttccc aacacataac tctagtatat 1440
aactatgcac caaatatgga taaacactgg attatgcagt acacaggacc aatgctgccc 1500
atccacatgg aattttacaaa cattctacag cgcaaaaggc tccagacttt gatgtcagt 1560
gatgattctg tggagagggt gtataacatg ctcgtggaga cgggggagct ggagaatact 1620
tacatcattt acaccgccga ccatgggttac catattgggc agtttggact ggtcaagggg 1680
aaatccatgc catatgactt tgatattcgt gtgccttttt ttattcgtgg tccaagtgt 1740
gaaccaggat caatagtccc acagatcggt ctcaacattg acttggcccc cagcatcctg 1800
gatattgctg ggctcgacac acctcctgat gtggacggca agtctgtcct caaacttctg 1860
gaccagaaa agccaggtaa caggtttcga acaacaaga aggccaaaat ttggcgtgat 1920
acattcctag tggaaagagg caaatttcta cgtaagaagg aagaatccag caagaatat 1980
caacagtcaa atcacttgcc caaatatgaa cgggtcaaag aactatgcca gcaggccagg 2040
taccagacag cctgtgaaca accggggcag aagtggcaat gcattgagga tacatctggc 2100
aagcttcgaa ttcacaagt taaaggaccc agtgacctgc tcacagtccg gcagagcacg 2160
cggaacctct acgctcgcg cttccatgac aaagacaaag agtgcagttg tagggagtct 2220
ggttaccgtg ccagcagaag ccaaagaaa agtcaacggc aattcttgag aaaccagggg 2280

```



```

actccaaagt acaagcccag atttgtccat actcggcaga cacgttcctt gtccgtcgaa 2340
tttgaagggtg aaatatatga cataaatctg gaagaagaag aagaattgca agtggttgc 2400
ccaagaaaca ttgctaagcg tcatgatgaa ggccacaagg ggccaagaga tctccaggct 2460
tccagtgggtg gcaacagggg caggatgctg gcagatagca gcaacgccgt gggccccacct 2520
accactgtcc gagtgacaca caagtgtttt attcttccca atgactctat ccattgtgag 2580
agagaactgt accaatcggc cagagcgtgg aaggaccata aggcatacat tgacaaagag 2640
attgaagctc tgcaagataa aattaagaat ttaagagaag tgagaggaca tctgaagaga 2700
aggaagcctg aggaatgtag ctgcagtaaa caaagctatt acaataaaga gaaaggtgta 2760
aaaaagcaag agaaattaaa gagccatctt caccattca aggaggctgc tcaggaagta 2820
gatagcaaac tgcaactttt caaggagAAC aaccgtagga ggaagaagga gaggaaggag 2880
aagagacggc agaggaaggg ggaagagtgc agcctgcctg gcctcacttg cttcacgcat 2940
gacaacaacc actggcagac agccccgttc tggaacctgg gatctttctg tgcttgacg 3000
agttctaaca ataacaccta ctggtgtttg cgtacagtta atgagacgca taattttctt 3060
ttctgtgagt ttgctactgg ctttttgga ttttttgata tgaatacaga tcttatcag 3120
ctcacaataa cagtgcacac ggtagaacga ggcattttga atcagctaca cgtacaacta 3180
atggagctca gaagctgtca aggatataag cagtgcacac caagacctaa gaatcttgat 3240
gttggaataa aagatggagg aagctatgac ctacacagag gacagttatg ggatggatgg 3300
gaagggtaat cagccccgtc tcaactgcaga catcaactgg caaggcctag aggagctaca 3360
cagtgtgaat gaaaacatct atgagtacag acaaaactac agacttagtc tgggtggactg 3420
gactaattac ttgaaggatt tagatagagt atttgcactg ctgaagagtc actatgagca 3480
aaataaaaca aataagactc aaactgtcca aagtgcaggg ttcttggttg tctctgctg 3540
gcacgctgtg tcaatggaga tggcctctgc tgactcagat gaagacccaa ggcataaggt 3600
tggaagaaaca cctcatttga ccttgccagc tgaccttcaa accctgcatt tgaaccgacc 3660
aacattaagt ccagagagta aacttgaatg gaataacgac attccagaag ttaatcattt 3720
gaattctgaa cactggagaa aaaccgaaaa atggacgggg catgaagaga ctaatcatct 3780
ggaaaccgat ttcagtgccg atggcatgac agagctagag ctcgggccca gccccaggct 3840
gcagccatt cgcaggcacc cgaaagaact tccccagtat ggtggtcctg gaaaggacat 3900
ttttgaagat caactatata ttctgtgca ttccgatgga atttcagttc atcagatggt 3960
caccatggcc accgcagAAC accgaagtaa ttccagcata gcggggaaga tgttgacca 4020
ggtggagaag aatcacgaaa aggagaagtc acagcaccta gaaggcagcg cctcctcttc 4080
actctcctct gattagatga aactgttacc ttaccctaaa cacagtattt ctttttaact 4140
tttttatttg taaactaata aaggtaatca cagccacca cttccaagc taccctgggt 4200
acctttgtgc agtagaagct agtgagcatg tgagcaagcg gtgtgcacac ggagactcat 4260
cgttataatt tactatctgc caagagtaga aagaaaggct ggggatattt ggggttgctt 4320
ggttttgatt ttttgcttgt ttgtttgttt tgtactaaaa cagtattatc ttttgaatat 4380
cgtagggaca taagtatata catgttatcc aatcaagatg gctagaatgg tgcctttctg 4440
agtgtctaaa acttgacacc cctggtaaat ctttcaacac acttccactg cctgcgtaat 4500
gaagttttga ttcattttta accactggaa tttttcaatg ccgtcatttt cagttagatg 4560
attttgcact ttgagattaa aatgccatgt ctatttgatt agtcttattt ttttattttt 4620
acaggcttat cagtctcact gttggctgtc attgtgacaa agtcaaataa accccaagg 4680
acgacacaca gtatggatca catattgttt gacattaagc ttttgccaga aaatgttgca 4740
tgtgttttac ctgcacttgc taaaatcgat tagcagaaag gcatggctaa taatgttggt 4800
ggtgaaaata aataaataag taaacaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4860
aaaaaaaaaa aaaaaaaaaa caaaaaaagc tgccgccaca gttagatgaa gaagcatgag 4920
gatccgagng ggtcgctctt ttgagtgtgt agggagtcgc g 4961

```

<210> 75

<211> 3298

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 1336022CB1

<220>

<221> unsure

<222> 76

<223> a, t, c, g, or other

<400> 75

```

ataaagttgg ccaacgggtg agtccccat tgtgaaaccc aagctttggg cattggaaat 60
ctcatttgag aaacantgat gagcagtagg cataacatca atagtttact gggcacatac 120

```

tggctagtat	tgccattttt	tagtttccta	ttgattttta	cttttttatt	ttttaatttt	180
tttttctttt	tttttttatt	taattttatt	gaggagtgcc	agacaatctt	tttattgttc	240
actgaaaaat	gcaggtctgc	aaagagtcaa	ttgcattgta	tattgaatgc	aaggctctgat	300
attgcaagta	tatatgacat	ggtataacat	ataaaatatt	acatatttta	cacagtgaca	360
gtacccgcct	cttctaaaca	ctaaaattta	atagaatgaa	gtaaaaagcc	tattaaataa	420
gaaacaaaca	ctgcaatcat	aaacaaaatg	cactaagcaa	aatactttaa	aattgttggtg	480
tgtgacatct	atcttgtcta	cctggagtta	ggaatgggtg	agcccagaaa	aaaactgagc	540
taacaaaaac	agccaacacc	ctccagggaac	tcttccctgg	aatcaggccg	ctgggtgccc	600
tcacaggctg	cccccatctg	atgaccacgc	caggcttcag	ggctgcctga	gggcccacac	660
tgaatgctaa	cgctgtctg	gggctgggct	gagggtgaatg	ggaggagtac	caggtaggaa	720
gaagcttctc	tcccaactgc	ttattctgca	tccatgctat	ggatatgaac	acctagtgtt	780
tctgggtgtga	atctgccact	ctccacaaac	attgaaatgt	gattgagaag	aaaaatgttaa	840
cgctgggaac	aaaagccatt	ttctacgcag	cttgcagata	tcatacttaa	ttcctttcat	900
ccatgttttc	attaagatgc	tctaaaagag	gctgtgctga	ttatacagcc	agtcaggcaa	960
gcaggatag	aggtaacaaa	atttctccag	ttataactag	taattcaa	gaatcaaaaa	1020
taaatatagc	ttatactaga	gagcaacaca	gtctctatta	ttcacatggg	aatatattca	1080
agaattgcaa	caatattcac	aattctctaa	ggccaacata	tttccctcca	agcagaaaa	1140
cctgaatgga	agggagtgtg	taacatcaat	aactcctgtt	tgtctcacat	aatatcaggg	1200
ctcagaaagc	tcattataga	taacaaatgc	aaaaataaaa	aaaaaaagca	tttcaatata	1260
tgtacagcaa	cagaggaaga	ctaggccctc	cctctgggccc	tccttggcat	aaggcaaaagc	1320
atggctatct	gcgacaacag	tgtctaagct	gactctggaa	aggctcatgt	tttacgataa	1380
atcagaatta	caaagttcct	tcattcacag	gtaacaaagg	attattaaat	aatttcctga	1440
actaagacaa	cagtctcacc	atcaaagaaa	attttcttct	catctgagtt	tttataaatg	1500
agattttaca	ataccagct	tttttagtta	atatgaaacc	atcggcgctac	agagaacatc	1560
ctcaaagctc	aagatgagta	aagctaatat	tctcagattt	tccacactct	acagagcagc	1620
aaatatttta	tcagaatatc	catttaacac	agttttaagt	ccataataac	caactatgta	1680
tttttcttat	tcaaagaaga	ccatgttttg	ataagaataa	tcccagtttg	atgagtatcc	1740
tattccctct	gatgttgtac	aataaatgtc	actgttgagt	ctctatatcc	ttgactcttc	1800
cagtcaaagg	aaaatcctgt	ataaagctgg	aaaactgcaa	actccagcgg	aagcactagt	1860
gttagtgctg	aagtaaactt	taggggaata	accacagtcg	cttatcagac	tgtttacagc	1920
attgatagaa	gccctcctct	tcttttatct	ctattatctc	taaacattct	aatattttta	1980
acatgggggt	tcactgatat	cgataaagg	ctagttcatg	agtaaggttt	tggtctctgg	2040
aatgatgtgt	ggctagccca	aaacactggc	tatgttagct	ttaaacaatga	tcattcttct	2100
tcccaccaca	atctgtcaaa	cccgcccatc	cgctctgcca	ctgtcactcc	cccagcccct	2160
caggattgct	gctcctctct	tctgaatgtt	ctactgcatg	tatgggtatc	actgagacat	2220
cactttaatg	tataggteat	ttgcatttca	tgccaccagc	tccccaaaag	acatcaactg	2280
ggaacaatct	ctaataagta	aggcacgatt	tgccagaatt	ttccctttat	aaaggaaatc	2340
aaaggacctg	tctggaagct	gaaagaacac	cttgggtatt	cctgtaaagg	gaacgctgta	2400
tctgtacca	aaacccagag	acactgacta	gatctaaaga	gcataatggat	tatgcaatag	2460
tgtgagatct	ggcagtttaa	gatctatttc	caatattaca	ttaagccatt	tggttgaaag	2520
tggtgttagt	atcacaggag	caaataaaaa	gtgtcaatga	cacccttttg	gctcatttcc	2580
ttgtgcttct	ggttataact	ccaccagctc	cagtgaagcc	tgtcccaggg	cacatcactc	2640
agctccccgc	acagctgctc	agagagaaaa	caatgcactt	cacctcaact	agtcagcaa	2700
ctggcactca	gatgggtgaac	gctgcagcca	acggccttgg	ggcagagccc	atggaaagct	2760
ttaaacaggc	ttatagacat	tgcataaaaa	ttcctgattt	taaaataacct	tcccaaggaa	2820
gtcacaaaaac	aatcattttt	agttaataaa	atgtaattct	tactcaaaaa	caatgtaatg	2880
cctcagaaag	caaaaacctg	aatggaagca	cttcagaaaa	atacggctta	atcagtatct	2940
tactttggct	aagttaccac	agtcagcgta	atttattatt	gcatttagaa	gcccaccta	3000
ccgtgttctc	ctaagatggg	gaagtacacg	caaacagtga	gcaggggtga	tagcccttta	3060
gccttttgtg	acttcagaac	ttaggaagat	aaacgataac	caggatttga	cattcagata	3120
cttacttgct	cagcacgagt	tgagttagga	aacagctaaa	gaccagattc	tcattccactt	3180
aaagtggtct	caggaagctc	ctatttggat	accaaaccag	tgggtctcaa	catggctgca	3240
catcagaatc	acctgggaag	cttttaaaac	tcaccacgta	gccaccattg	ccaatacc	3298

&lt;210&gt; 76

&lt;211&gt; 833

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7473674CB1

&lt;400&gt; 76

```

cacggcggtct gctggcggcc gcggagacgc agagtcttga gcagcgcggc aggcaccatg 60
ttcctgactg cgctcctctg gcgcggccgc attccccggc gtcagtggat cgggaagcac 120
cggcggccgc ggttcgtgtc gttgcgcgcc aagcagaaca tgatccgccc cctggagatc 180
gaggcgagga accattactg gctgagcatg ccctacatga cccgggagca ggagcgcggc 240
cacgcgcgcg tgcgcaggag ggaggccttc gaggccataa aggcggccgc cacttccaag 300
ttccccccgc atagattcat tgcggaccag ctgcaccatc tcaatgtcac caagaaatgg 360
tcctaatect gagtcgtcac ccttggattt tatggatcac ggagctgacc atctttacct 420
ggtcctggaa ctgaaaaact gtagcttgtg tgaaaatgag cctttggacc agtctttatt 480
aaaacaaaca aacatgagta gtctgcatat cgaatatcta gagctctaaa ccccccaata 540
cttaaaagtc taattgtctg cctgtggttt cattagtctg ataggaagat agggatttcc 600
tcagtcacag atgatatttt gaaggaaaag tgcaataaag ccacaatgat ttgaggtctt 660
tgcttaagta tgagatactt gatgggggct ttatcatgca acattagttt gcttacctta 720
agaattggcc aaaaatgaaa gaaaatatga gcttttcagt taaacatact cctaaaaaca 780
ttttccggga ttttactact aaaattggac atttaagcga agtaaaagag gcc 833

```

&lt;210&gt; 77

&lt;211&gt; 920

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7475846CB1

&lt;400&gt; 77

```

cacaccaacc ggcccaggcc ccgctggctt cagcctggtc acgacctcca cggcgcgggcg 60
ctgtctgctc ttggccgggc catcctgtcc tggccctccc tgcccacggc tccgggaagg 120
gctctgccaa gagacttctc cattcccacc cgtcactttt gaggatccgg ctccccggtc 180
cctcctgtgc cagccagcat gtgccatggc tccccgacct tgtgccagcc cgtgtgtgcc 240
atggctcccc accccgtgcc agcccatgtg tgccatggct cccaaccct gtgccagccc 300
gtgtgggcca tggcccacc gaaccctgtc caaccgcgt gtgccatggg ttcaccggac 360
ccgctgccag ccgcgtgccc ccatggcttc cctgatccca tgccagccc cgtgtgcgcc 420
atggctcccc cgaccccggt ccagcccggc tgcgtcatga cccaccgcg tgtgccat 480
ggcttccccg accccatgcc agcccgctg cgccatgggt ccaactgacc tgtgccagcc 540
agtgggggct gatctgctgc ccagctgtgg gcgtcctgaa cccaggccg tgggacaact 600
tgggcccggga gagaagatct ttcactttga tttggggcat aggtggagge tccccctacg 660
gccctgtgtg gaggaccctg ttctctgggt gctgtgatgc cactggcctc gaccctggct 720
cctcatgtgt gtgcctgccc ggctgaggt ccgtgcagag gaaggaagaa agaggaagta 780
agacgtgagc tcggccagcc ccgggtgcag gaggagtggg cgaagcgggc gacagggtcc 840
catgccttcc ccttccctcc tcaggccggc cctccatctc tcccaccaca ccaggcgcc 900
gctgggtgat ggcggggtca

```

&lt;210&gt; 78

&lt;211&gt; 964

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7475860CB1

&lt;400&gt; 78

```

cccatggcgc ctggttgtct agacgagccg gtgtaccgac tccccgggtg atgataacta 60
tgagtgcata agcaagacgc gcactttcct gtcacctcgc cctttcaccc cctcgggaag 120
atcaaagtgt cttagccca agcatggcct tgagcgactt catcacacct gcatgaggtc 180
acaagcagtc actagcgcag gggggtgagg cgagttcttg gtgaaagaaa cggggcgggg 240
ctgaggccaa ggaagaggtt ggactgtgtc tgcgatgtc ctgacgcaga aggttttgaa 300
ccttttttca cctcgtctga aatggctgcc tcccagtgtc tctgtgtctc aaaatttctc 360
ttccagagac agaacctcgc ctgtttcctc acaaaccac actgtggcag ccttgtaaat 420
gcagatggcc atggtgaagt gtggacagat tggataata tgtccaagtt tttccagtat 480
ggatggcgat gcaccactaa tgagaatacc tattcaaac gtaccctgat gggcaactgg 540
aaccaggaaa gatatgacct gaggaatata gtgcagccca aacccttgcc tttccagttt 600

```

```

ggacactact ttgaaacaac atatgataca agctacaaca acaaaatgcc actttcaaca 660
catagattta agcgagagcc tcaactgttc ccaggacatc aacctgaact ggatcctccc 720
cgatacaaat gcacagaaaa gtcaacttac atgaatagct attcaaagcc ttaaattggg 780
catcactcag gatgtgtata agatcttaat attgactagt ttcacatcca ggtttctaag 840
aaatgataag atacttcact tttccagagt gaaatgtagg agggagcaca ttctaagtac 900
agctaaaaat ttagctcact gtaacacagt ttcactctct gaataaataa agcaaaaaac 960
acag

```

<210> 79  
 <211> 701  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7950941CB1

```

<400> 79
ccgctactgc gctatgagaa accagcacaa ttatagcctt gagcattcta agcattttac 60
agtttacaaa taatcctttt atgatgttaa cagcttctct atttattaag tgcctagcat 120
gtgtcagacc tcatgttagg tattgtgtag gctttacctc acttaattct ggaagggtaca 180
tattctaatt ctttccattt tatagatgag gaaacaggct cagagagact aagttattgg 240
ataaggctac acagctcata aatgggtggag ctggatctca aaccagagc ctctggttgt 300
aagtcctgaa ttttaggaagc atctggaagc cctatcaagg gtgtagacac caagagtcac 360
cagaatggct ccaaatccag ctcgtttgca cagtcatttg gatttagtga gtccatccgt 420
accaaggtct ctgggctttc aacttcctat aggcaggaag cagtcaagaa atgtgctaag 480
ccaccaagat gggcatattc tccaatgttc ctttaggcca gacaggagga tgaaaaggaa 540
ggctgagagc ccagagaaca atcaactgag gtgccatctc ccatgccagg gtgggggaccc 600
agccatgttg ccagcagat ttcagaattg ctgaggacca gtgactgccg atgccctttc 660
cagcacatgg cggccgtaca agtgatgcga gctcgtagca g
701

```

<210> 80  
 <211> 1742  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7485334CB1

```

<400> 80
tcgttgaaat agttgggtgg ctaagagagg ggtctccacg tcgggggacgc ggaggggacc 60
tgccggagttg gcgtccgaac gcatagagcg gggccaccgc cgccgctcca ccattacctc 120
cccaggcggc aaggaggagc tgggtggcggt cgctcccgg ctgtggcagc ggcgggcgcg 180
cgcctgtctg gcggccgtcg gcgtgctctt ggccatggca ctgggggctgc tgatcgcggt 240
gcctctgctg ctgcaggcgg cgcccccccg agcggtcac tacgagatgc tgggcacctg 300
ccgcatgata tgtgacctat acagcgtcgc tcccgcaggg ggaccgcgcy gcgccaaggc 360
tccaccgccc ggaccagta ccgctgccct ggaagttatg caggacctca gcgccaacce 420
cccgcctccg tttatccagg gaccaaaggg tgatccgggg cgaccaggca agccagggcc 480
tcgggggtcct cctggagagc cagggcctcc tgggcccagg ggtcccccg gagagaaagg 540
agactcgggc aggccagggc taccggact gcagttgaca accagcgcg ccggtggcgt 600
tggagtgggtg agtggcgga cggggggcg tggcgacacg gagggagaag tgaccagtgc 660
gctgagcgcc gccttcagcg gtccaagat cgccttctac gtgggactca agagccccca 720
cgaaggctac gaggtgctca agttcgacga cgtggtcacc aacctcggca atcactatga 780
ccccaccacg ggcaagttca gctgccaggt gcgcggcatc tacttcttca cctaccacat 840
cctcatgcgc ggcgggcgac gcaccagcat gtgggcccgg ctctgcaaga acgggcagg 900
ccggggccagc gccattgcac aggacgcga gtatgtgaag ctggatggcg ggaagggtca 1020
ggtgctgcac ttggattcag gggacgaagt ctcgggcttt cttctgtacc cggattagg 1080
cggaggcaat aataacaagt acagcacgtt ctcgggcttt cttctgtacc cggattagg 1140
gcgcgggggg tgcgaggcgg ggtggctgca ggccgcccgc tctcggccc ggcgcggtc 1200
cttggaacaag gccactctcg attcataaca cttcctgaca tctccttttg aaaagacaaa 1260
tccctgcgct ctcctgccc cgctcctggc ctcagtgcgt ctgcgaccca ccagctcag 1320
ggctgtgctc ctgggtctcca tccccatccc ggcaaggag gaagggacgc ccgagccctt

```

gaggcgccgg cacagacttt gcaaacctga ttagcctgga caggcagggc cgggagcctg 1380  
ccctcctcag acagcctcct cccagtgcct agaagcggag ggctccgggc cctggccagg 1440  
gaggtaggcc agagggagcg cgggcttctt ggggcgtcct tctttgtgac ccgaaatact 1500  
tgtgcagatt tccctgtcca tcagccaaaa cccacccac agcagaattc cagcaaacag 1560  
aaaattcacc tctccacacc gcattccctc ctgactcaga ctaccgcga tgcattaaat 1620  
tatgttttta gaaaaaaaaa agaacaaaaa aaaagcaaaa aaaaaaagga aagggaaaca 1680  
caaataccga gagacaaggc ggtgccagaa aaaaaaaaag gggggggggc ctctttttatt 1740  
ta 1742

<210> 81

<211> 2295

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 7220001CB1

<400> 81

ggaaggatat ggatcagtgt tttctttttt gaagctactg ttaccactcc tggaaaagtt 60  
cttcagggaat aagtgcaggt aagaatgaca agggattagg actggcttcc tcttataaat 120  
aataaaatcc aaagagaagt gacttgagtc tccagggttta aagaagagca actagaagtc 180  
gtccaaacac ctgcatctca taaggagaag aaaagtcac ctggatcttg tttctggact 240  
gagatggatg gagaggccac agtgaagcct ggagaacaaa aggaagtggg gagggagga 300  
agagaagtgg actactccag gctcattgct ggcactttac cacaatctca cgtcaccagc 360  
aggaggcgag gatggaaaat gcccctcttc ctactactgt gctgtctaca aggttcttct 420  
ttcgcccttc caaaaaaaag acccatccg agatggctgt gggagggtc tctcccctcc 480  
aggacccatc tccgggccat gggaacactc aggccttctc cgcccctctg ctggcgggag 540  
gagagctcct ttgcagctcc aaattcattg aagggtcaa ggctgggtgc aggggagcct 600  
ggaggagctg tcaccatcca gtgccattat gcccctcat ctgtcaacag gcaccagagg 660  
aagtactggt gccgtctggg gcccccaaga tggatctgcc agaccattgt gtccaccaac 720  
cagtatactc accatcgta tctgtaccgt gtggccctca cagactttcc acagagaggc 780  
ttgtttgtgg tgaggctgtc ccaactgtcc ccgatgaca tcggatgcta cctctgcggc 840  
attggaagtg aaaacaacat gctgtttctta agcatgaatc tgaccatctc tgcagggtcc 900  
gccagcacc tccccacagc cactccagct gctggggagc tcaccatgag atcctatgga 960  
acagcgtctc cagtggccaa cagatggacc ccaggaagcc acccagacct taggacaggg 1020  
gacagcatgg gacacatggt gcttccacat ccaggaacca gcaagactac agcttcagct 1080  
gagggaaagc gaaccccagg agcaaccagg ccagcagctc cagggacagg cagctgggca 1140  
gagggttctg tcaaaagcac tgctccgatt ccagagagtc caccttcaa gagcagaagc 1200  
atgtccaata caacagaagg tgttcggggg ggcaccagaa gctcgggtgac aaacagggct 1260  
agagccagca aggacaggag ggagatgaca actaccaagg ctgataggcc aaggggaggac 1320  
atagaggggg tcaggatagc tcttgatgca gccaaaaagg tcctaggaac cattggggca 1380  
ccagctctgg tctcagaaac tttggcctgg gaaatcctcc cacaagcaac gccagtttct 1440  
aagcaacaat ctcagggttc cattggagaa acaactccag ctgcaggcat gtggaccttg 1500  
ggaactccag ctgcagatgt gtggatcttg ggaactccag ctgcagatgt gtggaccagc 1560  
atggaggcag catctgggga aggaagcgt cgaggggacc tagatgctgc cactggagac 1620  
agaggtcccc aagcaacact gagccagacc ccggcagtag gaccctgggg accccctggc 1680  
aaggagtcc cgtgaagcg tacttttcca gaagatgaaa gcagctctcg gaccctggct 1740  
cctgtctcta ccatgctggc cctgtttatg cttatggctc tggttctatt gcaaaggaag 1800  
ctctggagaa ggaggacctc tcaggaggca gaaagggtca ccttaattca gatgacacat 1860  
tttctggaag tgaaccccca agcagaccag ctgccccatg tggaaagaaa gatgtccag 1920  
gatgactctc ttcctgctgg ggccagcctg actgccccag agagaaatcc aggacctga 1980  
gggacagaga gatgaactgc tcagttacca tgggagaagg accaagatca aaggccttca 2040  
ggacccacgc ctctttccat catccttctc ccacctgtgg gaagagaagc tgatgcagcc 2100  
ggtgtccac ccatggaaga aaggctggct gtccttgggc ccaagaaagt caagcattat 2160  
ccacgtccag aaggtgacaa gatgactcaa aggagacttc aagaacagtg tatgaaacac 2220  
tggaagaggt cacctaggaa aagcatgaaa tttccagggg atccactagt tctaggcgcc 2280  
gcccgcgtg gctcc 2295

<210> 82

<211> 911

<212> DNA

<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 5956275CB1

<400> 82  
gcccgtctctcc gctccccgggc ccccgccggg cagcgcgccc ccccggggag atggaacagc 60  
ggaaccggct cggtgccctc ggatacctgc cgcctctgct gctgcatgcc ctgctgctct 120  
tcgtggccga cgctgcatc acagaagtcc ccaaagatgt gacagtagcg gagggagacg 180  
acatcgaaat gccctgcgcg ttccggggcca gcgagaccac ctctgtattcg ctggagattc 240  
agtgggtgta cctcaaggag ccaccccggg agctgctgca cgagctggcg ctgagcgtgc 300  
cgggcccggc gagcaaggta acaataagg atgcaactaa aatcagcacc gtacgcgtcc 360  
agggcaatga catctcacac cggcttcggc tgtctgcccgt gcggtgcag gacgagggcg 420  
tgtacgagtg ccgctgtctc gactacagcg acgacgacac gcaggagcac aaggcccagg 480  
cgatgctgcg cgtgctctcg cgtctcgcg cgcacaacat gcaggccgccc gaggcctgtg 540  
cccacatcca gagcagcggc ccgctgcgccc acggcccagc cagcgccgccc aacgccaaca 600  
acgcccggcg cgcgagccgt accacctccg agcccggcgg cggcgacaag agcccggcgc 660  
ccgggagccc tcccgcgcgc atcgatcccg cagtccccga ggcgcgggca gcctcggcgg 720  
cccacacgccc caccaccaca gtgcgaggcag ctgctgctgc ctgctcagcg tcgcccgcct 780  
cgggacaggc ggtcctgctg cgccagaggc acggctcggg taagggacgt agctacacca 840  
cagaccact cttgtccctg ctctgttag ctctgcataa gttcctgcgc ctgctcttgg 900  
gacattgaca g 911

<210> 83  
<211> 1806  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 346472CB1

<400> 83  
aaggctctca ggtgtttgaa gagatagttc taagtaaag aaaacaagca caaaaataag 60  
cacaagcatg gatgctttac atgagattac aaagagggga gagaggattg atggctccac 120  
agtttagcct aagggttagac tgggaattgat ggggtctgag gcctgctggg cttctctaac 180  
aggtggtcac tgtggacccc gggccttctt ctctctgatt ttaggtcatc ctgagctctc 240  
agagagtaga ccattcttaga ggcaggacc caggtaccct gaaattacc tccagctccc 300  
taggtctgca gtcctgctct cactgacagt ttcctctga agagaagcat cttcccctgt 360  
cctgactcag tgtctcagtt ggagcgtgt ggtctcccct ttcctcgggt tggctctcga 420  
gaactacctg ggaggctgct gggactccc catctcttcc tgcacagag cctgtgctcc 480  
ccttttgggc tgtccctggg gtttgtttct ctgttgaggg aggacactcc cagctctctg 540  
cacctttgag cagcagagat gtggagacac tgtcactggt tccagcagct ggctcaggtt 600  
acagcccacc cacttctctg ggagattgtc ttagctcat tgcccatatg gtcccagggc 660  
aagggtgggc ctattgttgg tcttttaatt aaaatatgat gttttaata 720  
tcccaaggat ttgcatactt cccaccagga attaagagg ctgttttgtc atagttaaag 780  
cccattgtaa agtaattgat aaaaaaatcc agtgccatga gttgagaggg ttagaaaaat 840  
aatagaaatc tttgttaata ctactctta tattatttta aatctacata gtgaatcttc 900  
cctgcctttt acccaacgat tctttccata aagaaaaata attaggttaa tgcaactgta 960  
attcaaagca aatatggat tcatcttttt tcttttttct ggatgccttt tatgttttag 1020  
tttcttacaa agcaatttcc agcactcaga taaaccattt gagaggaaca gacttagaat 1080  
tccatattca cagaactgtg ggatttttaa accacaaagg aaacctagag atcctagaag 1140  
actgttctgt ggatgtggaa agttcaaata tcccccaaga ctgcacagct aattagaggc 1200  
agagctgggg taaggactca ggtttcctgg ctctcaactc aaatctcttt ccaatatcat 1260  
ttctcaaatt ctaagatcag aagttgaatg aaacgattcg cttcacattt cttgtggttt 1320  
gaaacatgag cttctcagaa tctccagctt cagtggggtc agggtaacgt attggctgct 1380  
actgatggag gccagacaga gccccggcca ggttggggtc aggaatgtcg ttaggagaga 1440  
gggattccac atttaagctg cagagagagc agtgatgat gatagaagca ggaaatgcc 1500  
gaggccccag gggctcgccc aggtgaggaa acaggtacct ccactcgctc cttttagggtg 1560  
ggcctgccct gctgccagg gctaagtggc ttcccacagt ggctcagaaa catcaccgta 1620  
acccagtggt ttggaggaga tcatttcaac taagaggaaa aaacttttct ttttttaagt 1680  
aaaaaggatc tatttgaaat ttgtcattcc aaactagact tggagacagt tgtaaatctc 1740  
actcttttta ttctgggagc actctgtgct tttctagctc attctgttaa aaagaaaaaa 1800  
aaaagg 1806

<210> 84  
<211> 603  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 643526CB1

<400> 84  
gtagtcagtt accttcattt ctccgatgttt tcagagggct aatgctttgt gcagggtatt 60  
tatttgtagc tgaattattg tccttggttt cacagaggat atattaaagt atttttttgt 120  
gttgaagttt gggctgtgat ccagtagatg gtgcttaagc atgctggctg ataggttagac 180  
tgttggtcag tcacatgact actctgtatt tgcccgcatt tgcagctgtg ctctctctct 240  
ctcagtgctc tgagagtgtg ggctcctttc ccactcaagt gctggctgca gatctgggtt 300  
tggcactcct ggacgtcata ctacagcccc ggggtaagct cagcctttat gttccctcca 360  
cagcatgggg gcaaacacga accttgacag tggcaatggc agagggcctt taatttgtct 420  
cttggggctc tatcccagag acatgcagaa ctgctgtcaa tcagagtgat tggccttgtg 480  
tggggtgggt gcattgtggg ctcaagcctg ggggcccatg gggaacacag actggcctct 540  
tcttacagca actgcagcat gctggaggtg tgagtaaggc actcaggggt ctctgtttct 600  
tcc 603

<210> 85  
<211> 1888  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1483418CB1

<400> 85  
atgaaaggtg tgtggaaagt ggccagcatt ctgcctgcag accctgagca ctgagtcaac 60  
aagggtttgt atgatgtctg tactagttag ctattgttat taacaagatc tccatggcag 120  
acaagctaaa gcatctgcct aagtcgtgag tctgtgggtt ggctcagggt aagctgagct 180  
tggctgggct tggcagcaag tctcagattt gatcaagtca tatctgtttt cttcttcggc 240  
tgggtgggcta cccagagcga gcacgcttct ctgctgggtc tgccaagaac acaggagggt 300  
gagtcctaaca agcacacctc aagcctctgc ttgggtcaca ttggcaatgt cctctcagcc 360  
catacaagtg gcaggaccag gccaatattt aaggagtggg aaagtgtagg ctgcccttgg 420  
tgaagctgtg acaagggagt gaccatccat cccctccaca ggagggtgag gaattgggac 480  
aggtaattcc agccaatacc tgggtccatg ttcttcacat tgtgagtata tgaatacggg 540  
gaaatagaac cccctcccca ccacctacac acatgctgtt ttacctcca agaataaggac 600  
atgtgcaaaa actaggcttc cagatcgtga ggctggggag cattctaggt accatttatt 660  
atctactcgc ttctcgtcag tagctgttgt aggtaccgcg atcagccctt gttcctaagg 720  
atgtgtggag cccaggctgt actctttgac attccacagg tcagctcgcg tgaaagccct 780  
ttgtggggag ttccaagcc agtgtcgttt ttgcactggt ttccacctca ctgtcagata 840  
ttgttctttt gatatgcagt ttccagccct ccaatgagtc cccagcacag atctcaggct 900  
tgtggctggg aataagccca gctgagcagg tggggatctc agccagcctc ccagactgca 960  
cctcttctat tgetctctc tgtccaccgt tccaggcagc agcactccca cctggctgct 1020  
ctgatgtcct cgctccctcc attcctgcct ctccacctgc acgcatccat tgtccacaca 1080  
cacacaacca gataatcttc tgaaaagtga actcagatga tgctccaac cctgtgggta 1140  
cccttcacac ttacaggcct aaccccagca ggacctgcca gctgcttcat gaacaagtc 1200  
ctgcctcatg ttaattgcaa gcactcattt cactctccct gtgtgccagg caatagcata 1260  
agcactggaa gaatacagat gttatagtga gtccataaat agccctggaa ggcaagagct 1320  
gtttcctttc catttctaca atgagtcaat tgcagatgat gctataacac ttccagtgtt 1380  
tctgacactt cctgaagct atacctgcta ccttcatggg ccgagcttgc ccttatgagg 1440  
cccacaggtg gcagtgggca gaggggacc cgtataacca cctcgctgct gttccactgt 1500  
ctgctcccggt gttctgacca cagctctggt ccggtttctc aagcctggg ttcattcaac 1560  
atcttctatc tagctcttca tgggtgctgct ccgctatgg ttccacagg cttcttctcg 1620  
caggtcagct ccttagagag gtctcccaga ttcccgttaa agcagccctg cagcctctgt 1680  
ctctctcagc cgcatacccc tgttgcttcc ttccacagcat gtctcaccat ctgcaaccat 1740  
ctttctgttt gtgctcttgt tgatttgctg cctccacact gtcagctcct tgggaacaga 1800  
gattgggtttg ttactgtgc atccctggtg cccagaacag ggcattggcat attgttggtg 1860

cataataaat atggtggaaa ctaaaaaa

1888

&lt;210&gt; 86

&lt;211&gt; 1576

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2683477CB1

&lt;400&gt; 86

aaacaaatta	gcaatgtaaa	gaacatttgt	tgagttcagt	gttggagaga	gcagttcaca	60
gagctgccga	gaagccacag	ctgacccttt	cctctctcct	cttggtagga	gccagctgca	120
aggagagggg	cctaaggtgg	tagaggggaat	ggctccctcc	tccacagctc	tgcattgcgtc	180
agcccccaaa	atagaaatgc	ggggaccaag	ttgtgatggc	agggaaacag	cagaaagagg	240
tgaggctgcc	tgtgctccct	gccctccagg	cctccacagg	ccaacctgtg	acctcactgg	300
tgggcctcca	gaagcaggga	aagacagagg	cccagcatca	cctttccatt	tccgcatttg	360
ttttctcttc	cttgctgggt	tgctattgact	ctgtgggtcac	tgttctccat	gtcagcacat	420
tcaaaattgc	tgactgtcaa	cactgaaggc	agcgtggctg	ctactgaaga	agccacaagg	480
aaaacagctt	tgggcaatgg	tggatgttct	tgggtgtgaca	catttctagc	tcccagcaca	540
ggccctttac	aaagaacagg	gctttgtggt	ttgggttagga	tggggagaaa	gaaagaggga	600
gggagagaga	gaaggaggct	tggcttgttg	agatcttttg	ttaaggaaat	aaatattggt	660
ttcctagaat	ttgaacagct	gaaatgggag	attggcagta	agcaaaatgt	gattgtaaca	720
atcaaaccga	cgtgcagaca	gaagttggga	gtcattaggg	caatgagggt	gttcttcacg	780
atcttgggag	gaaagaaaag	gtgaccaaaa	tgccatttag	taaccgatg	gcctcttcaa	840
gcccttcagg	ctggcccaga	gctgcaggca	aagctttgat	ggtgtgggtg	gtgctgttcc	900
cttgggcaga	gcttggctgg	aggactctta	gcagggtggc	cgcaagtctc	tggggcccct	960
acttggggac	ttacacagac	caggctgtat	gtctttgtag	cttgtcaaac	cacaactatt	1020
cacagaaggc	gtgtggttta	gaatctacca	cagtcaaacc	cgggagaatg	tgttaccag	1080
ttccagaaag	gttgctagtt	tgtgtgctgt	aatggaaaag	ggccaacttt	atcttttctt	1140
ataagaacta	caatgggcaa	tcaggatgga	ggcttgatct	acaggactct	ctactgaggg	1200
tctcctgtga	aaaagggtgat	aggagggagg	tgctggttgt	attaatcaga	gtcctctaga	1260
gacagaacaa	atagaggata	aagagagaaat	tcattttaaa	gaactggctt	atttgactgt	1320
gggggctggc	aagtccaaaa	tctgtagacc	agactaccag	gctagaaatt	caagtgatct	1380
gcccgtttca	gcctcccaaa	gtgttaggat	tacagggtgtg	agccactgca	cgcactccaa	1440
tagcccttcg	taagctttaa	tttaattggc	ccatccttct	tcacccttg	gcccccttaa	1500
tactagactt	gtaccacaga	aattttctgt	gatctttttt	tgcttaccaa	actttctctg	1560
tactggagac	aaaaaa					1576

&lt;210&gt; 87

&lt;211&gt; 415

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 5580991CB1

&lt;400&gt; 87

ctttcccgaga	aagccttttg	aaaccctgtt	ttattatgaa	aatattcaaa	cattcacaaa	60
actagagaaa	atacctattc	ctagattcaa	cagtgatgaa	cataatgcc	tatttgcttc	120
aactttcatt	cttctctctc	cttttctccc	tccctttttc	tctctgccct	tcttctctct	180
ctctattgtt	ttttcttttg	gctgtggggg	tttatttttt	ttttgagacg	agtcttgctc	240
tgtcacccag	gctggaatgc	agtgggtgca	tctcggctca	ctgcaagctc	tgccctcccg	300
gttcattgcta	ttcttggggc	tcagcctgtg	gagtagctgg	ggctacaggt	gcccaccacc	360
acgcccgggc	aatttttggg	atttttagta	gagatggggg	ttcactgtat	tggcc	415

&lt;210&gt; 88

&lt;211&gt; 762

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens



<220>  
<221> misc\_feature  
<223> Incyte ID No: 5605931CB1

<400> 88  
ttttgtggtt gttttttgtg ttttttgatt tgtgttgttg tgccaggtaa taatgtagtg 60  
ccatgcctgt gttcttcttc ttctttttta aatagagaca gattcttgct atggctatgg 120  
ctacagtctt aoccaaagctc ttcttgaagc cccggcctca agtgatcctc ccacctcaac 180  
ctcccaaggt gctgagatta caggcatgag ccatcatgct cagccttcaa ccagatttct 240  
tgacagcacc tttatatgtg attcttacac caatcttcat tgttaacccc attttaaaga 300  
agcagaaacc aaggcccagg gaggtacat aacttgccag agctctcaga gccaaaaagc 360  
aacagatgtg gaattcaaac ccgggcattc agatgccaa gttccctgca ctcccactct 420  
cccaaactgc ctgagcctga gcaagcccaa cctgaagcct tctcctgga gtccaaagtc 480  
cagccaggaa tgtgacatgg gctccccagc cctccagatg tgtgtgctca cactttgtct 540  
ggacttggtc ctccctggcc ttcgaaagtt ctgcctcag atgtcccat tagtcacagt 600  
ctgtctgaga gccctcggat tagctggatg ggagcagacg caactttgtg gtgggtcatca 660  
ggttgtccca ttcatcagct caggcctgag cctgctggag tgtgggtcgt gccagaaaca 720  
ataactctta ggaaagaaaa ccagcctcgt gccgaattct tg 762

<210> 89  
<211> 654  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 6975241CB1

<400> 89  
gtttaaacgc gcacccccag ggcgagacct gggttaaagc aaaagggcta acaagggcaa 60  
cccagaatcc ggcagggctg cagtgaaccc tgcccagcac actctctgct gtgaccttgc 120  
cgtgattcta ggttcacacc ttcccatcag tgctgtgagc tcttggtgct gtcagcgatg 180  
gtttcatctg ttccatttag acaaagtcag gttctagttt tgtgcctgtg tctgtgccta 240  
gaacagaaat tggatccagg tgtgatttgc aagcaagaga tctcagaga aatggggtatg 300  
tgggaggaca ctggagtgc cagatctagt tgtactgaag tcaataaaaa tccagctggg 360  
tcttcctgga tgggaatcca gcagaccagg gctcacaca gcggaagagc tacataact 420  
ggtgcctgtg attggctgca atggagccca ttgagggcaa gagatccagc tgccataaaa 480  
caggagaagc tacaggtagg gagccgattc taatgcatgg agaactacta ggctagagca 540  
aaggccagc ttctccatgc aacgcgcaga agggttccta gagcttccag cattatgctg 600  
agagaatctt ttctaaggct ccggtgtgtc atagcagaaa accagaggca aaac 654

<210> 90  
<211> 505  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 6988529CB1

<400> 90  
tttagccaaa gaatcaaagc attggtgaag aattgaagat tggaagttac acattttttg 60  
ctaaggggaag taatagagaa agaaaaatat tcttgagggt aaatcttcag tttggattaa 120  
tgtgaacatg aagaaacctg tagaaacctt catttctcaa gacaaagctc aaattcaagg 180  
ttgtgaggaa tgcagtcact gcttttggtt ggggcagttg tgacagtgat tgcagaaact 240  
gagattgcaa agcccgtatt atataaagaa tgtgcaagtg ccatagaaga cactgcaagg 300  
attgggtgct ggagcagtgc tggacctgcc gtcacacca gagtgagca gagagaatct 360  
cctcctttgc catcactaac ccagcacttg actttgtccc actcctaaaa gatcttgaa 420  
ggaattaaag gtgcttggtg gtttctcatt ccagcataac cttacttggc ctgacctga 480  
accatgcatt agtctttgga caggg 505

<210> 91  
<211> 841

<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 6996808CB1

<400> 91  
ttcaatgagg gtgactcatt ttgtgactga gacaaacatt cacatgtgga ccatgttaaa 60  
aataaataaaa taaatggcag gtgctgcttg gctaaattct gtgttttaaat gctgtaattt 120  
cctgccgtaa gggttcacgt cttgtataat gtctactcag cctctgtaat cactagccca 180  
atatatctaa tgcacctcaa ttcattgcta ctcatttctt gattatttat tttagttcct 240  
aagggtcttt tgaattgtaa caatgtgatt tcaacaggga agccaaggaa tggatggggg 300  
ggagccacag ttagccggaa aacaaagagt gacacataag ctgtagcaaa aggacatgtt 360  
ctgtgctttt ctgttttttg cattttccca agatgttttg tgcattgtgt ttggtaaagt 420  
tgtcttagtt atgtttatatt tattatgtat atgttcagta ttggagctct tttctcaag 480  
tggaagatgc tttgaaagca ctctcttcat tgtggcccat gtatcaaatac taatctcaaa 540  
aattttacaa gtctactctc tcaggagaat tctgtttatt tattgtacag atatgctatg 600  
tactaggcac tgtgctatgg caaactaaag caggcggtgg ccactatctt ctgaaactt 660  
tttcagatct gtttggggag ttacgtacta ccacaaatac atcctcacgc aaattgaatt 720  
gcaaattacg cacatcacta cgtaggtaag taaggatctc taaatatgca ttagttacct 780  
gagtggtaga tgagggtggc gaagaagcct agttggtgcc gcaaaggga atcgttggga 840  
a 841

<210> 92  
<211> 1367  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 7472689CB1

<400> 92  
agacattaga ttggagattc agacctctcc gtgatgtgac atttgagcca agatgatagg 60  
gagccagacc tgggaaattt ggggacagga aacagcaagt gccgaggctc tgagggtggca 120  
gaggaacaga aagaagtga atttggtctg gggtagttaa ctcatgtctg agagtcataa 180  
tggggacatg tgctaagatg cgtgagaaag tgctctggaa actctttagt gctgcataaa 240  
gggttaccgt tgcttttgct gatgttcgtg tgatgtatgt ttcaggacct ctagtgacct 300  
tgaacaagcc acagggtcta ccagtgcacg gtatgggccg gggatgtggg aagggaatgc 360  
agcgggaagg ggtttcgtga ctgagggagg gaaaagtga ggcatgaagc tttggcctct 420  
tgtaattttt ctttcttact ttccaggaaa accaggagag ctgacgttgt tctcagtgct 480  
gccagagctg agccagtcct tagggctcag ggagcaggag cttcagggtg tccgagcatc 540  
tgggaaagaa agctctgggc ttgtactcct ctccagctgt cccagacag ctagtgcct 600  
ccagaagtac ttcacccatg cacggagagc ccaaaggccc acagccacct actgtgctgt 660  
cactgatggg atcccgctg cttctgaggg gaagatccag gctgccctga aactggaaca 720  
cattgatggg gtcaatctca cagttccagt gaaggcccca tcccgaagg acatcctgga 780  
aggtgtcaag aagactctca gtcactttcg tgtggtagcc acaggctctg gctgtgccct 840  
ggtccagctg cagccactga cagtgttctc cagtcaacta cagggtgcaca tgggtactaca 900  
gctctgccct gtgcttgggg accacatgta ctctgccctg gtgggcactg tctgtggcca 960  
gogatttctg ctgccagctg agaacaacaa gcccaaaga caggctcctg atgaagccct 1020  
cctcagacgc ctccacctga cccctccca ggctgccag ctgcccttgc acctccacct 1080  
acatcggctc cttctcccag gcaccagggc cagggacacc cctgttgagc tctggcacc 1140  
actgccccct tatttctcca ggaccctaca gtgcctgggg ctccgcttac aatagtcctc 1200  
cctctgttcc tgacccctc acacacactg gaaagtggg gtgggggctc tgcagtcaga 1260  
caaacctaa atcacatcct ggacaggcca cttgcttgc gtgtggcatt gggcaagtaa 1320  
ctttacctct ctggacttgt gataataaaa gttcctacct caaaaaa 1367

<210> 93  
<211> 4595  
<212> DNA  
<213> Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 876751CB1

&lt;400&gt; 93

```
gagggggctc cgggcgccgc gcagcagacc tgctccggcc gcgcgcctcg ccgctgtcct 60
ccgggagcgg cagcagtagc ccgggcgccg agggctgggg gttcctcgag actctcagag 120
gggcgccctc catcgccgcc caccacccca acctgttcct cgcgcgccac tgcgctgcgc 180
cccaggaccc gctgcccac atggattttc tcttgccgct ggtgctggta tctcgtctct 240
acctgcaggg ggcgcgcgag ttcgacggga gtaggtggcc caggcaaata gtgtcatcga 300
ttggcctatg tcgttatggt gggaggattg actgctgctg gggctgggct cgccagtctt 360
ggggacagtg tcagctgtg tgccaaccac gatgcaaaca tggatgaatg atcggggcaa 420
acaagtgcaa gtgtcatcct ggttatgctg gaaacacctg taatcaagat ctaaagtgt 480
gtggcctgaa gccccggccc tgtaagcaca ggtgcatgaa cacttacggc agctacaagt 540
gctactgtct caacggatat atgctcatgc cggatgggtc ctgctcaagt gccctgacct 600
gctccatggc aaactgtcag tatggctgtg atgttggtta aggacaaata cggtgccagt 660
gcccaccccc tggcctgcag ctggctcctg atgggaggac ctgtgtagat gttgatgaat 720
gtgctacagg aagagcctcc tgccctagat ttaggcaatg tgtcaacact tttgggagct 780
acatctgcaa gtgtcataaa ggcttcgac tcagtatat tggaggcaaa tatcaatgtc 840
atgacataga cgaatgtctc cttggctcagt atcagtgcag cagctttgct cgatgttata 900
acgtacgtgg gtcctacaag tgcaaatgta aagaaggata ccagggtgat ggactgactt 960
gtgtgtatat cccaaaagtt atgattgaac cttcagggtc aattcatgta ccaaagggaa 1020
atggtaccat tttaaagggt gacacaggaa ataataattg gattcctgat gttggaagta 1080
cttggctggc tccgaagaca ccatatattc ctctatcat taccaacagg cctacttcta 1140
agccaacaac aagacctaca ccaaagccaa caccaattcc tactccacca ccaccaccac 1200
ccctgccaac agagctcaga acacctctac cactacaac ccagaaaagg ccaaccaccg 1260
gactgacaac tatagacca gctgccagta cacctccagg agggattaca gttgacaaca 1320
gggtacagac agaccctcag aaaccagag gagatgtgtt cattccacgg caaccttcaa 1380
atgacttggt tgaaatatatt gaaatagaaa gaggagtcag tgcagacgat gaagcaaagg 1440
atgatccagg tgttctggta cacagttgta attttgacca tggactttgt ggatggatca 1500
gggagaaaaga caatgacttg cactgggaac caatcaggga ccagcagggt ggacaatatc 1560
tgacagtgtc ggcagccaaa gcccaggggg gaaaagctgc acgcttggtg acgctctcgc 1620
gccgcctcat gcattcaggg gacctgtgcc tgtcattcag gcacaagggt acggggctgc 1680
actctggcac actccagggt tttgtgagaa aacacgggtg ccacggagca gccctgtggg 1740
gaagaaatgg tggccatggc tggaggcaaa cacagatcac ctgcgagggt gctgacatca 1800
agagcgtcgt cttcaaagggt gaaaaaaggc cactgctctg aagaacgcta acaactccag aactaacaat 1920
atgtgagctt gaaaaaaggc cactgctctg caattctcat cttctctcct cttctcctct 1980
gaactcctat gttgctctat cctcttttct cttctctcct cttctcctct 1980
ttatcaggcc taggagaaga gtgggtcagt gggtcagaag gaagtctatt tggtagacca 2040
ggtttttctg gcctgctttt gtgcaatccc aatgaacagt gataccctcc ttgaaataca 2100
ggggcatcgc agacacatca aagccatctg tgggtgttgc ctccatcctt gtgtctcttt 2160
caggaaggca ttcagcatgc gtgagccata ccactctcca tcttgattac aagggtgctc 2220
ttgtagcaaa ttatgagagt gagttacggg agcagttttt aaaagaaatc tttgcagatg 2280
gctatgatgt tatgtgttcg gtgtgtgacc atgtagtaga ttgacttccc ttgagatag 2340
atgtacaatg tgcttgtgaa attgacttac cctcttcact taagttagtt ctggcctgac 2400
ctgaactctg acttttactg ccattcactt tataaaataa ggggtgtgta catatcaaga 2460
tacatttatt tttatctgtt ttttttttct ctgttaaaga caattatgta gagtgggcac 2520
gtaatccctc cttagtagta ttgtgttttg tgtaaattgt ctattgatat taagtattta 2580
catgttccaa atatttacag actctagtgt caaggtaaag ggcagcttgt gatctcaaaa 2640
aaatacatgg tgaaatgtca tccagttcca tgaccttata ttggcagcag taggaaattg 2700
gcagaagtgt tgggttgttg taacggagtg atgaattttt ttttaattgg cttgagtttg 2760
atctctgcaa aggataggaa accttttaga agacaagaaa ctgcagttaa tttagaactg 2820
tactgtttc aagttacact ttaaaaccac agcttttacc atcataacat ggctctggta 2880
atatgtagga agctttataa aagtttttgt tgattcagaa aaaggatcct gttgcagagt 2940
gagagggaag atagggggaa actccatttg aacagatttt cacacaacgt tttaaattga 3000
tataagttta ggcagttgta gttcataact tatgttgctc atgttgtgct gtgtcaggat 3060
gggataggaa gcaagtccca tgcttagagg catgggatgt gttggaacgg gatttacaca 3120
cactggagga gcagggcaag ttggaattct aagatccatg aacccccaac tgtatttctc 3180
ccctgcatat tttaccaata tattaaaaaa caatgtaact tttaaaaggc atcattcctg 3240
aggtttgtct taatttctga ttaagtaatc agaataattt ctgctgtttt tgccaggaat 3300
cacaaagatg attaaagggt tggaaaaaaa gatctatgat ggaaaattaa aggaactggg 3360
attattgagc ctggagaaga gaagactgag gggcaaacca ttgatggttt tcaagtatat 3420
gaagggttgg cacagagagg gtggcgacca gctgttctcc atatgcacta agaatagaac 3480
```

```
aagaggaaac tggccttagac tagagtataa gggagcattt cttggcaggg gccattgtta 3540
gaataacttca taaaaaaaga agtgtgaaaa tctcagtatc tctctctctt tctaaaaaat 3600
tagataaaaaa tttgtctatt taagatgggt aaagatgttc ttaccaaggg aaaagtaaca 3660
aattatagaa tttcccaaaa gatgttttga tcctactagt agtatgcagt gaaaatcttt 3720
agaactaaat aatttggaca aggccttaatt taggcatttc cctcttgacc tcctaattga 3780
gagggattga aaggggaaga gcccaccaa tgctgagctc actgaaatat ctctccctta 3840
tggcaatcct agcagtatta aagaaaaaag gaaactattt attccaaatg agagtatgat 3900
ggacagatat tttagtatct cagtaatgtc ctagtgtggc ggtgggtttt aatgtttctt 3960
catgttaaag gtataagcct ttcatttgtt caatggatga tgtttcagat tttttttttt 4020
ttaagagatc cttcaaggaa cacagttcag agagattttc atcgggtgca ttctctctgc 4080
ttcgtgtgtg acaagttatc ttggctgctg agaaagagtg ccctgcccc caccggcaga 4140
cctttccttc acctcatcag tatgattcag ttctctttat caattggact ctcccagggt 4200
ccacagaaca gtaatatttt ttgaacaata ggtacaatag aaggtcttct gtcatttaac 4260
ctggtaaagg cagggctgga gggggaaaaa aaatcattaa gcctttgagt aacggcagaa 4320
tatatggctg tagatccatt tttaatgggt catttccttt atggtcatat aactgcacag 4380
ctgaagatga aaggggaaaa taaatgaaaa ttttactttt cgatgccaat gatacattgc 4440
actaaactga tggagaagt tatccaaagt actgtataac atcttgttta ttatttaagt 4500
ttttctaaaa taaaaaatgt tagtgggttt ccaaatggcc taataaaaac aattatttgt 4560
aaataaaaaa actgttagta ataaaaaaa aaaaa 4595
```

&lt;210&gt; 94

&lt;211&gt; 4759

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2512510CB1

&lt;400&gt; 94

```
atggcgcgcc cgggtccgggg aggggtcggg gcccgcgcc gctcgccttg ccttctcctt 60
ctctggctgc ttttgcttcg gctggagccg gtgaccgccc cgcccgcccc gcggcgcccc 120
tgcgcggccg cctgcacttg cgctggggac tcgctggact gcggtgggag cgggctgggt 180
gcgttgcccc gggacctgcc ctctgggac cggagcctaa acctgagtta caacaaactc 240
tctgagattg acctgctgg ttttgaggac ttgccgaacc tacaggaagt gtacctcaat 300
aataatgagt tgacagcggg accatccctg ggcgctgctt catcacatgt cgtctctctc 360
tttctgcagc acaacaagat tcgcagcgtg gaggggagcc agctgaaggc ctacctcttc 420
ttagaagtgt tagatctgag ttgaacaac atcacggaag tgccgaacac ctgctttcca 480
cacggaccgc ctataaagga gctcaacctg gcaggcaatc ggattggcac cctggagttg 540
ggagcatttg atggtctgtc acggtcgtg ctaactcttc gcctgagcaa aaacaggatc 600
acctcagctt ctgtaagagc attcaagcta cccagggtga cacaactgga cctcaatcgg 660
aacaggattc ggctgataga gggcctcacc ttccaggggg tcaacagctt ggaggtgctg 720
aagcttcagc gaaacaacat cagcaaaactg ctggtagaag tgaacagcgg ctgcgtctac 840
atgcatgtgc tgcacctgga gtacaacagc agcaacaatt ccatcgctcg cattcacgcg 900
ggcctcacgg cctgcatca gctccacctc gaagctgcat gaggttggtc tgtccttcaa caacctgaca 960
aagggctgga gcttctgcca gaagctgcat gagggtgctc gtgtcctgag tctcagccc 1020
cggctggagc agggagagcct ggccgagctg agcagcctga gtgtcctgag gctcagcttg 1080
aattccatca gccacattgc ggagggtgct ttcaaggagc tcaggagcct gcgagcttg 1140
gatctggacc ataacgagat ttccgggcaca atagaggaca cgagcggcgc cttctcaggg 1200
ctcgacagcc tcagcaagct gactctggtt ggaaacaaga tcaagtctgt ggctaagaga 1260
gcattctcgg ggcctggaagg cctggagcac ctgaaccttg gagggaatgc gatcagatct 1320
gtccagtttg atgcctttgt gaagatgaag aatcttaaag agctccatat cagcagcgac 1380
agcttctctg gtgactgcca gctgaagtgg ctgccccctg ggctaatttg caggatgctg 1440
caggcctttg tgacagccac ctgtgcccac ccagaatcac tgaagggtca gagcattttc 1500
tctgtgccac cagagagttt cgtgtgcgat gacttctctg agccacagat catcaccag 1560
ccagaaacca ccatggctat ggtgggcaag gacatccggg ttacatgctc agcagccagc 1620
agcagcagct ccccatgac ctttgctctg aagaaagaca atgaagtctt gaccaatgca 1680
gacatggaga actttgtcca cgtccacgag caggacgggg aagtgatgga gtacaccacc 1740
atcctgcacc tccgtcaggt cactttcggg cacgagggcc gctaccaatg tgtcatcacc 1800
aaccactttg gctccaccta ttcacataag gccaggctca ccgtgaatgt gttgccatca 1860
ttcaccaaaa cgccccacga cataaccatc cggaccacca ccatggcccc cctcgaatgt 1920
gctgccacag gtcacccaaa cctcagatg gctggcaga aggatggagg caggatttc 1980
cccgtgccc gtgagcgagc catgcatgtc atgcccggat acgacgtgtt tttcatcact
```

```

gatgtgaaaa tagatgacgc aggggtttac agctgtactg ctcagaactc agccggttct 2040
atttcagcta atgccacct gactgtccta gagaccccat ccttggtggt ccccttggaa 2100
gaccgtgtgg tatctgtggg agaaacagtg gccctccaat gcaaagccac ggggaacctt 2160
ccgccccgca tcacctggtt caagggggac cgcccgctga gcctcactga gcggcaccac 2220
ctgacctctg acaaccagct cctggtggtt cagaacgtgg tggcagagga tgcgggcccga 2280
tatacctgtg agatgtccaa caccctgggc acggagcgag ctcacagcca gctgagcgctc 2340
ctgcccgcag caggctgcag gaaggatggg accacggtag gcatcttcac cattgtctgc 2400
gtgagcagca tcgtcctgac gtcactggtc tgggtgtgca tcatctacca gaccggaag 2460
aagagtgaag agtacagtgt caccaacaca gatgaaaccg tcgtgccacc agatgtttcca 2520
agctacctct cttctcaggg gaccctttct gaccgacaag aaaccgtggt caggaccgag 2580
ggtggccctc aggccaatgg gcacattgag agcaatggtg tgtgtccaag agatgcaagc 2640
cactttccag agcccgcac tcacagcggt gcctgcaggc agccaaagct ctgtgctggg 2700
tctgcgtatc acaaagagcc gtggaaagcg atggagaaag ctgaagggac acctgggcca 2760
cataagatgg aacacggtgg ccgggtcgta tgcagtgcac gcaacaccga agtggactgt 2820
tactccaggg gacaagcctt ccacccccag cctgtgtcca gagacagcgc acagccaagt 2880
gcgccaaatg gcccgagcc ggggtgggagt gaccaagagc attctccaca tcaccagtgc 2940
agcaggactg ccgctgggtc ctgccccgag tgccaagggt cgctctacce cagtaaccac 3000
gatagaatgc tgacggctgt gaagaaaaag ccaatggcat ctctagatgg gaaaggggat 3060
tcttcctgga ctttagcaag gttgtatcac ccggactcca cagagctaca gcctgcatct 3120
tcattaaact caggcagtc agagcgcgcg gaagcccagt acttgcttgt ttccaatggc 3180
cacctcccca aagcatgtga cgccagtcct gagtccacgc cactgacagg acagctcccc 3240
gggaaacaga ggggtgccct gctgttggca ccaaaaagct aggttttgtc tacctcagtt 3300
cttgtcatac caatctctac gggaaagaga ggtaggagag gctgcgagga agcttgggtt 3360
caagcgtcac tcactctgac atagttgtaa ctcccatgtg gagtatcagt cgctcacagg 3420
acttggatct gaagcacagt aaacgcaaga ggggatttgt gtacaaaagg caaaaaaagt 3480
atttgatatt attgtacata agagttttca gagatttcat atatatcttt tacagaggct 3540
attttaatat tttagtcatg gtttaacagaa aaaaattata caattttgac aatattatatt 3600
ttcgtatcag gttgctgttt aattttggag ggggtgggga aatagtctct gtgccttaac 3660
gcatggctgg aatttataga ggctacaacc acatttgctt acaggagttt ttggtgcggg 3720
gtgggaagga tgggaaggcct tggatttata ttgcacttca tagaccctca ggctgctgtg 3780
cggtgggact ccacatgcgc cggaaggagc ttcaggtgag cactgctcat gtgtggatgc 3840
ccctgcaaca ggcttcctg tctgtagagc caggggtgca agtgccatcc acacttgacg 3900
tgaattggct ttctctttag gtttaagtcc tgtctgtctg taaggcgtag aatctgtccg 3960
tctgtaaggc gtagaatgag ggttgttaat ccatcacaag caaaagggtc gaacagttaa 4020
acactgcctt tctcctcct cttattttat gataaaagca aatgtggcct tctcagtatc 4080
attcgattgc tatttgagac ttttaaatta aggtaaaggc tgctggtgtt ggtacctgtg 4140
gatttttcta tactgatgtt ttcgttttgc caatataatg agtattacat tggccttggg 4200
ggacgaaag gaggaagttc tgacttttca gggctacctt atttctacta aggaccaga 4260
gcaggcctgt ccattgccatt ccttcgcaca gatgaaactg agctgggact ggaagagaca 4320
gcccttgacc tgggttctgg gtataatttg cacttttgag actggtagct aacctctta 4380
tgagtccaa tgtgtcattt agtaaaactt aaatagaaac aaggctcctc aaatgttctc 4440
ttggccaaaa gctgaaggga gttactgaga aaatagttaa caattactgt caggtgtcat 4500
cactgttcaa aaggtaagca catttagaat tttgttcttg acagttaact gactaatctt 4560
acttcacaa aatatgtgaa tttgctgctt ctgagaggga atgtgaaaga gggagtatta 4620
cttttatgta caaagttatt tatttataga aattttggtt cagtgtacat tgaaaacct 4680
gtaaaatatt gaagtgtcta acaaattggc ttgaagtgtc ttaataaag gttcatttat 4740
aaatgtcaaa aaaaaaaaaa 4759

```

&lt;210&gt; 95

&lt;211&gt; 3203

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7486326CB1

&lt;400&gt; 95

```

ccctcctccc cagctgtccc gttcgcgtca tgccgagcct cccggccccg ccggccccgc 60
tgctgtcctt cgggctgctg ctgctcggtt cccggccggc ccggggcgcc ggcccagagc 120
ccccgtgct gccatccgt tctgagaagg agccgctgcc cgttcgggga gcggcaggct 180
gcaccttcgg cgggaagggt tatgccttgg acgagacgtg gcacccggac ctagggggagc 240
cattcggggt gatgcgctgc gtgctgtgct cctgcgaggc gcctcagtgg ggtcgccgta 300

```

```

ccagggggccc tggcaggggtc agctgcaaga acatcaaacc agagtgccca accccggcct 360
gtgggcagcc ggcgcagctg ccgggacact gctgccagac ctgccccag gagcgagca 420
gttcggagcg gcagccgagc ggccgtgtcct tcgagtatcc gcgggacccg gagcatcgca 480
gttatagcga ccgcgggggag ccaggcgctg aggagcgggc ccgtgggtgac ggccacacgg 540
acttcgtggc gctgctgaca gggccgaggt cgcaggcggt ggccagagcc cgagtctcgc 600
tgctgcgctc tagcctccgc ttctctatct cctacaggcg gctggaccgc cctaccagga 660
tccgcttctc agactccaat ggcagtgtcc tgtttagca cctgcagcc cccacccaag 720
atggcctggt ctgtgggggtg tggcgggagc tgccctcggt gtctctgcgg ctccctaggg 780
cagaacagct gcatgtggca cttgtgacac tctctaccc ttcaggggag gtctgggggc 840
ctctcatccg gcaccgggccc ctggctgcag agaccttcag tgccatcctg actctagaag 900
gccccccaca gcaggggcgta gggggcatca ccctgtcac tctcagtgc acagaggact 960
ccttgcatctt tttgctgctc ttccgagggc tgctggaacc caggagtggg ggactaacc 1020
aggttccctt gaggtccag attctacacc aggggcagct actgcgagaa cttcaggcca 1080
atgtctcagc ccaggaacca ggctttgctg aggtgtgccc caacctgaca gtcaggga 1140
tggactggct ggtgctgggg gagctgcaga tggccctgga gtgggcaggc aggccagggc 1200
tgcgcatcag tggacacatt gctgccagga agagctgcga cgtcctgcaa agtgtccctt 1260
gtggggctga tgccctgatc ccagtccaga cgggtgctgc cggctcagcc agcctcacgc 1320
tgctaggaaa tggctccctg atctatcagg ccgtgggtat ctgccctggg ctgggtgccc 1380
gaggggctca tatgtctgtg cagaatgagc tcttctgaa tgtgggcacc aaggacttcc 1440
cagacggaga gcttcggggg cactggtgtg ccctgcccta ctgtgggcat agcggccgcc 1500
atgacacgct gcccggtgcc ctagcaggag ccctggtgct acccctgtg aagagccaag 1560
cagcagggca cgcctggctt tcttgata cccactgtca cctgcactat gaagtgtgc 1620
tggctgggct tgggtggctca gaacaaggca ctgtcactgc ccacctcctt gggcctcctg 1680
gaacgccagg gcctcgggcg ctgctgaagg gattctatgg ctcagaggcc cagggtgtgg 1740
tgaaggacct ggagccggaa ctgctgcggc acctggcaaa aggcattggc tccctgtga 1800
tcaccaccaa ggttagcccc agaggggagc tccgagggca ggtgcacata gccaaccaat 1860
gtgaggttgg cggactgcgc ctggaggcgg ccggggccga gggggtgcgg gcgctggggg 1920
ctccggatcc agcctctgct gcgcgcctg tgggtgcctg tctccggcc ctagcgccc 1980
ccaaacctgg tggctcctgg cggccccgag accccaacac atgcttcttc gaggggcagc 2040
agcgcacca cggggctcgc tggggcggcc actacgaccc gctctgctca ctctgcacct 2100
gccagagacg aacgggtgatc tgtgacccgg tgggtgtgcc accgcccagc tgcccacacc 2160
cggtgcaggc tcccgaccag tgcctccctg tttgccctga gaaacaagat gtcagagact 2220
tgccagggct gcccaaggagc cgggacccag gagagggtg ctattttgat ggtgaccgga 2280
gctggcgggc agcgggtacg cgggtggcacc ccgttgtgcc cccctttggc ttaattaagt 2340
gtgctgtctg cacctgcaag gggggcactg gagaggtgca ctgtgagaag gtgcagtgtc 2400
cccggctggc ctgtgccag cctgtgcgtg tcaacccac cgactgctgc aaacagtgtc 2460
cagtgggggtc gggggccccc cccagctgg gggaccccat gcaggctgat gggccccggg 2520
gctgccgttt tgctgggcag tgggtcccg agagtcagag ctggcacccc tcagtgcctc 2580
cgtttggaga gatgagctgt atcacctgca gatgtggggc aggggtgcct cactgtgagc 2640
gggatgactg ttactgcca ctgtcctgtg gctcggggaa ggagagtcca tgctgttccc 2700
gctgcacggc ccaccggcgg ccagccccag agaccagaac tgatccagag ctggagaaag 2760
aagccgaagg ctcttaggga gcagccagag ggccaagtga ccaagaggat ggggcctgag 2820
ctggggaagg ggtggcatcg aggaccttct tgcattctcc tgtgggaagc ccagtgcctt 2880
tgctcctctg tcctgcctct actcccacc ccactacct tgggaaccac agctccaca 2940
gggggagagg cagctggccc agaccaggt caccgacct ccaagtccg cctgccacc 3000
ctcgccctct gtccctggaag cccaccctt tctcctctgt acataatgtc actggcttgt 3060
tgggattttt aatttatctt cactcagcac caagggcccc cgacactcca ctctgtctgc 3120
ccctgagctg agcagagtca ttattggaga gttttgtatt tattaacaa tttctttttc 3180
agtcaaaaaa aaaaaaaaaa aaa 3203

```

&lt;210&gt; 96

&lt;211&gt; 1681

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1221545CB1

&lt;400&gt; 96

```

tttaataatc aaacactctg ataacctatg acaaaggatt agtatacaat taaatgaagt 60
aatgcataaa gaggttgggc atataataaa gacttacctc atgtgagcta aaaccactat 120
caaggcaggt tctaggacca agagcacaag tgaccttaaa tgccactgaa agctcccttg 180

```

gaggtgttct caccaggagg attagagaag ccaaaacaac cagggtgaata tctctgtcaa 240  
tgatagacaa cttggctata acaggaaaag attctagaat ttatgggatt atggaacaat 300  
aaatagagtt aacttttagaa aggagattta caaaataagt agcgggtatg gatattgcta 360  
gtccgtagct gattaaggct ctgattaagt gaattgcccc aagtctcaga gcagacatag 420  
gcctagtcca aacttagage tcatattact tgcagtggat gtttgttctc ttggctgtcc 480  
agcaggccac cttttccttc aggacactgc tctcccactg catttcacat gtgactcggt 540  
tggggctgcc aaaatatatt tggttatctt tttttttttt ttcagtaagc taatacaaaa 600  
ttactgtttc ttcaataatt actttgcatt tattgtcatc atttcactcc caaatgtatc 660  
aaaattaaag tttaagaggg aggaaaaaag gataagtaga agatcctgat taccttttag 720  
tcatagatag tttcattatg ttatctttta ggggtctggaa tacactgacc agtgtattag 780  
acaaaatttt atgagaattg gttaaagata tagggaagaa atgtatttgg aaaaaaaca 840  
accaaaccac accaaaacaa aacaaaaacc attattttga gagtaaatac ttggggggaa 900  
gaagcttgca agccaccgca atatggcctg gactctggcg tgtgtgtgcg tgctggggag 960  
tatcttggtg ttggactctg gcatgtgtgt gcgtgctggg gagtgtcttg atgggtgatgt 1020  
tgtgtcattg cttcattttt ggcactcagt cactactcaa gaaaaccaga ttgaaaattt 1080  
ggaatctgtg cttcagtggg ttgaaactgg cctccagtca ctaaggaaaa aatcaaaaaca 1140  
aaacacacaa gaatttagag agaataattt tctgccaaaa aataattttt cctttatgct 1200  
atttcttatt tgggtcaata ctccaatgga aaaaatagat agattgggtca agagtccaat 1260  
ataattttct gtgacatttg aactaaagta actcataaaa acttaaaccac aggaaattgt 1320  
atcctctcct gctgatgtat gtgtgactat ttgtctctct taaaagaaaa aagtagaaga 1380  
agagaattaa ttgaatggta ttttgtttta cttgagatgt taaactaatg tcaactaat 1440  
ttattttatt aataaatatt tattaagcac ctactacatg ctgggtactac aagccagga 1500  
tagatgcttg ggatgtatta atgagcaaaa caaaaccagt agcaaaactg tacacgcaga 1560  
taagggtttc aaatattgtt ggcattgggc aggtgcaatg gatcacacct gtaatttctt 1620  
ttctcttctt ttctctgctt ctttcttctc ttcttctctc ccttcttctt ttccttccgc 1680  
c 1681

<210> 97

<211> 1207

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 124737CB1

<400> 97

ccaaataact ctggaaccct ttaactgtcc agcagggaga tccaactacc ttgaaaccca 60  
ccaagctgga aaagcccccac agggcagcca cctctgtggc acaatcccca gctgaaacct 120  
tgcccttcca gactgtccct gtcaaagatg cccaggcaat gtgaataaag cccatcatgg 180  
accctctaaa ccagactgcc caccagcagg gtaccatctg gcagccacat ggagggcagt 240  
ccaccttccct gagtccacaa ttcaaagtct gatctaatac agaaacactc tcacagacac 300  
agtgtgtagt ctgggcaccc cacagtgcac ttaagtggac acataaaatt aaccatccca 360  
gggagtgaag tgggtaaggg aagatgggccc acagtgggtg tgtctccatg cctaccccca 420  
ctgtgggcag ctgctggagc acacgcctca aagtcttccc tgaggggagag ggagctgagg 480  
tgttttatatc ccagctctgt caggcattgg ctgaatgtcc acactcctgg atcaccgcct 540  
ctcactctca tgatgtccca tggctctcat ttcaaaagt agctctgggt acatggggag 600  
catcagtcac accctgggtc agtacctcag ctgtctctca catgacatcc tcattatcca 660  
cactgcaaag ccaaccatcc ctatgatggc ttcatgtgtg atcatgactt agtgggtcaa 720  
gagtttggaa gtggctcagc tgggcgggtc ttctgtctcca tgtggctgcc agatgggtacc 780  
ctgctgggtg gcagtctggt ctgaggggtc catgatggct ttactcacat gcctgggcac 840  
ttgacagggg cagctggaag gcaagggtca gctgggactg tccacagagc tccctccctgt 900  
ggcctttcca gcatgggtgg ctgagggtag ctggacttcc tgcatgacag ctgagggtc 960  
ccagagctac tgtcccaaga gatagaaggt ggaaactgcc agtctcttag gctaggacca 1020  
gaaaccagca cccctgcacc cacagccttt tggtagtgtat gaaataaaca taagatttat 1080  
cattttaatc attcgttaagt gggattaaat acatttaca tattgtgtaa ccacgggcac 1140  
tgtctatatc taaaactttt tcatcatctg caataaaaaa tctgtatgca ttaaaaaaaa 1200  
aaaaaaa 1207

<210> 98

<211> 1544

<212> DNA

<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1510784CB1

<400> 98  
tgaagggtgtg tggagcattc ctgacctttg tccccagccc tgccctctcc tgtcctgcag 60  
cctgcatctt tgctgagttc tcagggcctt ccagctagaa gttctgccat ctgttaaagt 120  
cgtatgtttc ctctcccgtt gcctgtctgc ctccctctcg gagtgcacct acagagcact 180  
agccctccat tccctgccag ccacacccag gtctccctct ctgactccca cacctgcctc 240  
actgccagcc cagctaaggt tctcttcaaa tgtctgtttt ctgtctgcct ctgccattcc 300  
cagtgtgacc actcgtgctc agccgtatct cagcaggagg acaggtgccg gagcagctcg 360  
tgcagctaag cagccaactg cagaaacgtc aggtgggtgg tgcatctgca ggcatgctga 420  
agaagcagtt ccaggcatgg gccgccggca gagaggctgg ctgcagcacc cccaccacc 480  
atgcacatcg tgctttgctt ctctctagag tttagcgctt tggcaggcca tgccttgttt 540  
gtgttcacag cctgctgttc agggagcgaag acagaacgtc tgacttttgg atccgtaaca 600  
gggggtggggg ctccccagaa gagaagaggc tttgggtcct ggccagtgtc cactactcag 660  
tcaaacattc agaaacttac atgattcttc atcgtcccaa gcaagtctga cttgggccct 720  
ttattgaaac tctgtttcct ctcccgtgc ctgtctgcct cctctcggga gtgcacctac 780  
tgagcaccag cctctcattc cctgccagcc acaccagaa agcaatgggg cttctctggga 840  
aggcagaaat attctgtgac ctgggctatt cagaggggtg gcgtgagtcg tgtgtgccta 900  
gcaaacactc aggacatggc cgggtgagaca gaagggtgta gcttttctcc tagctaatat 960  
attaattaat tcattcatta tttattttga gacggagtct ctctgtcgcc caggctggac 1020  
tgcagttgag cagtcttgcc tcaactgcaac ctctgcctcc taggttcaag caattctccc 1080  
gcctcagcta cttgggaggc tggggcagaa gaatcgcttg aaccaggag gcggagggtg 1140  
cagtgaagcg agattgcgcc actgcactcc agcctgtgag acagagtgcg actccgtctc 1200  
aaaataaata atgaatgaat taattaatc attagctagg acaaaagctc agcccttctg 1260  
tctcaccggc catgtcctga gtgtttgcta ggcacacacg actcacgcca cccctctgaa 1320  
tagcccaggt cacagaatat ttctgccttc ccagaaagcc ccattgcttt ctgggtgtgg 1380  
ctggcaggga atggagggtt ggtgctcagt aggtgcactc cgagagggag gcagacaggc 1440  
agcgggagag gaaacagagt ttcaataaag ggcccaagtc agacttgctt gggacgatga 1500  
agaatcatgt aagtttctga atgtttgact gtagtagtggg gctg 1544

<210> 99  
<211> 1519  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1901257CB1

<400> 99  
aagcctccaa gggaggcatc caattcactt ttggttaaga tcaatacttt cttcttcac 60  
aatcccctcc tgattaattt ttttaattgg ctttcagaag aaagcaaata tcaaagttgc 120  
ttctaaacat cactgatttt gctagctcta tcaattcatt tttttctatc aagtttttaa 180  
gataaccttg tgtgactcag gccattcctt gtttgacgt taccatcaa tacaagtcag 240  
ccaagacgag tgtcgctaga gcttcagtg tctttcacat tctagccctc ttcaaccaca 300  
aattataaaa acgtggctct gctcaagcac acgttttaaa ttaaaccctt ttgtttttta 360  
catgaatttt ttaggtcttt ttttcagggt attattttct gagacagtcc aataaaaatt 420  
tattttaaaa tgtattttgt gtaatttgat gacagcctca aaaaaatcac ataattagga 480  
ttttattaca aaagtcaaca gttcagttt tgttctggaa gtggggaatg gaaggaggga 540  
aagaggagga gcaggggagag aaaagatgag gaacctggta actgcaaaaa acaattcaag 600  
cagttatatt tcaccatgta cagtctggaa agaagagttt tctagaatca aggaagaaaa 660  
taaaagctct gttagtttgc tctgcatctt gctatgcctt tctaattaaa tgattggaag 720  
gacttcatta ttgactcctg ctggccgaca tgacactaaa atgatatgca tctcaatctg 780  
catactccaa gccaaaaccc aacatgccat atgcattgca catgtccttc caaaggcttt 840  
gggtctggat cctccttccc accgtggcca acattgctt gtctcatca agaactggca 900  
gatccaagga gcatacccaa gatgacgcca cagcctacat gctctctcgg cacctacatg 960  
ctctctcggc acctacatgc tctctcggca gcctacatgc tctctcggca gcctacacgc 1020  
tctcttggca tgtacagcag gttcttcagc cgtgcccagg aggcctgggg ctccgtgggtc 1080  
tatctctgag ctgggttcta gacctccac cccacttcca ccactgcaac ttctgtttta 1140  
catgttggaa aggggcttct tataatatgc cacttaaaga aaaagactga atttttttta 1200  
aataaaaaat atactggcct atgtcattaa aatgaaatat atcccaataa agttgtaaa 1260



caaaaagcaa actcttttcaa atcttatttta ctgcaaaaaca ttttagaaac tttcctcaat 1320  
tgccaccgat tttccaaagc agacctgtga aagccagcaa tgaaaaattt aaggttatta 1380  
ctcatacctg gctcttttgg aagaaggctg gacattagct acttcattct gtttcagttt 1440  
gggaggtagt cttatactct gcaattaaaa tattgtcgac ttttaattcaa tcaatctact 1500  
aagtaataca gtagcttcc 1519

<210> 100

<211> 525

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2044370CB1

<400> 100

agagttcctt tttctaggct gattagggtta tacattgttg aagtatagtt tcgagttaga 60  
attgggtcatt ttatttttcag tgtttcacag aaatcgaaga agacagaaat ggcgcttctg 120  
tggtggatat ctacagtagc aatactgttg tttacttcga cgattttggg aacatacgtt 180  
gaagctgggtg ccgctaagtc taacgaagaa gagattgtga acaaaagcga atttggaaga 240  
tttccacgag ggtcgagaaa ggatgcatcg ggggtgccaca agccgggcta cctgtaccc 300  
cctcattctc gctgccctcc acctcccat gtgcagcgtc ctgctctat tctgcatgct 360  
tagtctaaca ccacaggtc cgtttatctt ttctgtcatt gatctcacca ggagcaaattc 420  
actagtgcgt gcttctgatt cacgtaacgt agtatgtaaa taaatgtcag tgatattatg 480  
aattggtaaa acatttctgt tatctaaata aaacagtga agttt 525

<210> 101

<211> 1062

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2820933CB1

<400> 101

agtcgagaat caatctagca tccactctta gatattctatt aaaaggaggt aagaaacata 60  
tgtccactca gacttatagg caaatcagtc gtagcaccat aattcactat agtgtgacaa 120  
caggaaacaa ttcaggtgcc cctcaactga caaatatata aacaaatgta ttatatccat 180  
gcaatgggat actgttcagc aattttacaa aggactagt atgcctgcaa caacatggat 240  
gatectcaaa atgagctaag tgaaagaaac cagacacaga agatacatat taaatgattc 300  
cattcacatg aaatttctag aaaaggcaaa actatagaag caaaagcagg tcagtgcctg 360  
ggaatggcat tgacagcaag tgggcacaag aaaatttggg ggtgataaaa atgttctaaa 420  
actagattgt gatgatagtt gcacaactgt gtacatttac taaggctcat caaactgtgc 480  
ccataaaatg ggtagatatt gtgttattta aatgacacct taattaaggt gtttaaaaga 540  
aaaattcccc tgagtttctc ctgcttgcat ttgaagtga aaccagctc ctcatgggg 600  
ggccacctcc cccggcagc tctttctgtc tctgcttcat ccatggggct ttctccagtt 660  
tctcacccca cctccttcc catgagtgt ccagcaggtg ctgttcctc tgctggcac 720  
gctttcttgc ttctccact ccttgagta actcagagtc ttcttcaact ctctacctca 780  
agtcacgtct cgcaggaagc ctctctggct ctgcccactg cagtcctacc tccctgccct 840  
tctccttggg gacactcatc actcctgaaa ctggtgactc ttctcctaag tacagttct 900  
ggctcatagt ggggtgctcaa taaatatttc ataaaataat gtctgaataa tcatctgaat 960  
tgttctgaga agcccatgat aaacaaacct gtttaacttt gtgtaaccag tgtatctgag 1020  
gcatgttttc acaagaacct cacttttttt tttaatgggt cc 1062

<210> 102

<211> 2155

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2902793CB1

&lt;400&gt; 102

```

gcctgaggag cccacgaagg ggctccttcg tgggtacttc gtgatgttac caggctgaac 60
agaggggaatc tacagcctgt gcttggcata ccctgcatgt ggtctgtgtc cagttgggct 120
ttgtgtcttc tctgtgccat ccatgtcctg tctctttcat gtgtcfaatg taattgtgtc 180
catgtctttc tgatccctcc accagccctg cctgccaggt tcacagaggg tctgaggaat 240
gaagaggcca tggaaggggc cacagccaca ctgcaatgtg agctgagcaa ggcagcccct 300
gtggagtggg ggaaggcct tgaggctctc agagatgggg acaaatacag cctgagacaa 360
gacggggctg tgtgtgagct gcagattcat ggctgggcta tggcagataa cgggggtgtac 420
tcatgtgtgt gtgggcagga gaggacctca gctacactca ctgtcagggg taaagatcct 480
atgtggccat gtgggcttgt ggcttggtgt atacacctct ctgtgtcacc acctctctgcc 540
tccaaatgtg gcacatctcc tgtggaaacc ctgtgattgt ctgtcctcta actggggccc 600
tctaggcttc tctgctcc ccttctcacc agtagatgtc ctctgtctca tgcattgtcc 660
tgtgtgttct ctgtgctcc tggggcattt gccttctcat tgtttatata ttgtctatct 720
agctatgggt ctttgggtgt ttgtgtggct ccttatttgt gtccatgttc tgtccgaaaa 780
atcctccaga cagtctgatg atatcagtga ctgtttggct cttctcagcc ctgctgccca 840
gattcataga ggatattgag aaccagaagg ccacagaagg ggctasagtc acattgcaat 900
gtaagctgag aaaggcgggc cccgtggagt ggagaaaagg gcccaacacc ctcaaagatg 960
gggacaggta cagcctgaag caggatggga ccagttgtga gctgcagatt cgtggcctgg 1020
tcatagcaga tcttgagaa tactgtgca tatgtgagca ggagaggacc tcggccacgc 1080
tactgtcag gggtaaagac cacatgtgac cactgtgtg acttctgtct tccccactt 1140
aaccacatg ttctgtgtc tccagtgct tctcagtgct gttgacattt tattcagtca 1200
ctcatctttg tggctccatc cacaatgca tgtgaggac ccacttggat ggcctaactt 1260
agggcctggg catacagacc ccaagggtga atagtgcagg gtcccaggt gtcaggacag 1320
ggagagcagc aggcaggtgt cagggtccag gagagctgtc tggggcccct gccatcttgc 1380
aaaggcctgt ggatgtccac cagctcttga gcctcaggca tggaggtcag gaaatgtatg 1440
tcttttgaca gacatagcga tggctcagga caggccctc ttcaggctgg tgagtgttct 1500
gattgatcct tgtggtcatt ccaagcttct aaaggagtat tgtcttcacc tgctcaggct 1560
aacgtataaa agcactgcag acagggtggc tcaaataaca cagatttatt ttctcagaag 1620
tatggggctg aaatctcaag atcaagatgt cagctcctct cttctttagt acagctgtct 1680
tctccccatg tcttcacatg gatgtccctc tgtgtgcatg tgtgtgtcct aacctctttt 1740
tataagagca ccagtcatac tggatttggg accacctaa caacctcatt ttctcttcca 1800
atgattgtct ttaaatacag tgatattctg acgtgtactg ggggttagga cttcaacata 1860
tgcatttttt ctgaggcaca attcagacca taacactcca ccatctggat tctcaaaatt 1920
catggccttc tcacgtgcaa aatatgttta cttcttccca acagtcccaa atcttagccc 1980
atttcaatat caactagtcc aaatcacctc taaatatcat ctgagtcaac tgtgaggagg 2040
ataaagtgtg acttatacag aggcaaaatt tttcttcacc tgtgaggctg tgaaatcaga 2100
caagttattt gttttgaagt cacaatgggt ggacaggcat aggatggaca ttctg 2155

```

&lt;210&gt; 103

&lt;211&gt; 1777

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7486536CB1

&lt;400&gt; 103

```

gcctggactg tgggttgggg gcagcctcag cctctccaac ctggcaccga ctgcccgtgg 60
cccttaggca cctgcttggg gtccctggagc cccttaaggc caccagcaaa tccataggaga 120
ccgagtcttg gcacgtgaac agagccagat ttcacactga gcagctgcag tcggagaaat 180
cagagaaagc gtcaccacag cccagattcc gaggggcttg ccagggaact tctcctcctg 240
ctccttggaa aggaagaccc cgaaagaccc ccaagccacc ggctcagacc tgcttctggg 300
ctgcccattg acttgccggc accgcccccc ggctgtcctc cacgtgcgcg ggcagataag 360
ggcagctgct gcccttgggg cacctgtctc ctcccgcagc ccagccactc ctccagggcc 420
agcccttccc tgactgagt accacctctg ctgccccgag gccatgtagg ccgtgcttag 480
gcctctgttg acacactgct ggggacggcg cctgagctct cagggggacg aggaacacca 540
cgatgccccg gggcttcacc tggctgcgct atcttgggat cttccttggc gtggccttgg 600
ggaatgagcc ttggagatg tggcccttga cgcagaatga ggagtgcact gtcacggggt 660
ttctgcggga caagctgcag tacaggagcc gacttcagta catgaaacac tacttcccca 720
tcaactacaa gatcagtgtg ccttacgagg ggggtgttcag aatcgccaac gtcaccaggc 780
tgcagagggc ccaggtgagc gagcgggagc tgcggtatct gtgggtcttg gtgagcctca 840
gtgccactga gtcggtgcag gacgtgctgc tcgagggcca cccatcctgg aagtacctgc 900

```

```

aggaggtgga gacgctgctg ctgaatgtcc agcagggcct cacggatgtg gaggtcagcc 960
ccaaggtgga atccgtgttg tccctcttga atgccccagg gccaaacctg aagctgggtgc 1020
ggcccaaagc cctgctggac aactgcttcc gggatcatgga gctgctgtac tgctcctgct 1080
gtaaacaaag ctccgtccta aactggcagg actgtgaggt gccaaagtcct cagtcttgca 1140
gccagagacc ctcatgtcag tatggcgcca cccagctgta cctccgccc ccgtgggtccc 1200
ccagctcccc gcctcactcc acgggctcgg tgaggccggt cagggcacag ggcgagggcc 1260
tcttgccctg agcacccctg atgggtgact cggatagggg cagccagacc agctcccaca 1320
ggagttcaac tgggtctgag acttcaaggg gtggtggtgg gagccccct tgggagagga 1380
cccctgggaa ggggtgtttt cctttgaggg ggattctgtg ccacagcagg gctcagcttc 1440
ctgccttcca tagctgtcat ggcctcacct ggagcggagg ggacctgggg acctgaaggt 1500
ggatggggac acagctcctg gcttctcctg gtgctgccct cactgtcccc ccgcctaaag 1560
gggggtactga gcctcctgtg gcccgcagca gtgagggcac agctgtgggt tgcaggggag 1620
acagccagca cggcgtggcc attctatgac cccccagcct ggcagactgg ggagctgggg 1680
gcagagggcg gtgccaagtg ccacatcttg ccatagtga tgctcttcca gtttcttttt 1740
tctattaaac accccacttc ctttgaaaaa aaaaaaa 1777

```

&lt;210&gt; 104

&lt;211&gt; 2587

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 8137305CB1

&lt;400&gt; 104

```

aggcggcgcc gggcccaagg cgtgaggcgc cgcccggtg tccccgcggc gcaggaggcg 60
gtggagcgca gagcgggcga gcgcgaaaaa tcactaccaa tataatggat tttatatatc 120
agattgcttt attctggata tcatggtaac aatacagaaa gtatacataa tttcccat 180
ctgcaagtag tcatgactgc tgaagaaaga aaaacttaaa gctacggcag aattatttta 240
tggaaattct gattttgttt ttaatttttg ataactttt actaaaggta tgaacacaca 300
aagagcttat tttgttaggc aaatacacat taataagaat gcctagaaga ggactgattc 360
ttcacacccg gacccactgg ttgctgttgg gccttgcttt gctctgcagt ttggtattat 420
ttatgtacct cctggaatgt gccccccaga ctgatggaaa tgcattctct cctgggtgtg 480
ttggggaaaa ttatggtaaa gagtattatc aagccctcct acaggaacaa gaagaacatt 540
atcagaccag ggcaaccagt ctgaaacgcc aaattgccca actaaaacaa gaattacaag 600
aaatgagtga gaagatgcgg tcaactgcaag aaagaaggaa tgtaggggct aatggcatag 660
gctatcacag caacaaagag caagcaccta gtgatcttt agagtttctt cattcccaaa 720
ttgacaaagc tgaagttagc ataggggcca aactaccag tgagtatggg gtcattccct 780
ttgaaagttt taccttaatg aaagtatttc aattggaaat gggctcact cgccatcctg 840
aagaaaagcc agttagaaaa gacaaacgag atgaattggt ggaagtattt gaagcgggct 900
tggaggtcat taataatcct gatgaagatg atgaacaaga agatgaggag ggtccccttg 960
gagagaaact gatatttaat gaaaatgact tcgtagaagg ttattatcgc actgagagag 1020
ataagggcac acagtatgaa ctctttttta agaaagcaga cttacggaa tatagacatg 1080
tgacctctt cgcctcttt ggacctctca tgaagtgaag gagtgagatg attgacatca 1140
ctagatcaat tattaatatc attgtgccac ttgctgaaag aactgaagca tttgtacaat 1200
ttatgcagaa cttcagggat gtttgtattc atcaagacaa gaagattcat ctacagtggt 1260
tgtatttttg taaagaagga ctgtctaaag tcaagtctat cctagaatct gtcaccagtg 1320
agtctaattt tcacaattac accttgggtc cattgaatga agaatttaat cgtggacgag 1380
gactaaatgt ggggtgcccga gcttgggaca agggagaggt cttgatgttt ttctgtgatg 1440
ttgatattta tttctcagcc gaattcctta acagctgccg gttaaatgct gagccaggta 1500
agaaggtgtt ttacctgtg gtgttcagtc tttacaatcc tgccattgtt tatgccaacc 1560
aggaagtgcc accacctgtg gagcagcagc tgggttcacaa aaaggattct ggcttttggc 1620
gagatttttg ctttggaatg acttgtcagt atcggttcaga tttcctgacc attggtggat 1680
ttgacatgga agtgaaaggt tgggggtggag aagatgttca tctttatcga aaatacttac 1740
atgggtgacct cattgtgatt cggactccgg ttctgtgtct tttccacctc tggcatgaaa 1800
agcgtgtgct tgatgagctg acccccgagc agtaccgcct gtgcattccag tctaaagcca 1860
tgaatgaggg ctctcactcc cacctgggaa tgctgtgtct cagggaggaa atagagacgc 1920
atcttcataa acaggcatac aggacaaaca gtgaagctgt tgggttgaaat cataattaat 1980
gcgttactgt atgaaccaca aaacagcact atttatttag ccttacttct acttccagat 2040
gcagtgcctc ttttgagaaa gacatgttta tttttcatgt tctttctgac attactttag 2100
caattcaact tgatgtgaga agaaaaaaca aatgtttcaa cacaaaatct ctgttttgtg 2160
agaatactgc actatggaat aattgacaaa ttgaaatctc atatttgtcc caaaagtgtg 2220

```

```

tttgagtttag ttctacctgg tgcccatggt ctgatttgtg tgtgggattg catggtgtcc 2280
tgatgcatct aggtggagcg gatggaaatg tgctggagcc actggtgggt gagaagcaag 2340
aacgatactt accagaagga gattggagcg ttagtgagca ataggtatgt aggggaatagg 2400
gtatctatca aacgtgcaca gaacactgaa ataccagcct tacttggaat tgatagcttg 2460
aaagaatcaa ttaagccaca tgaagtagaa ggatactaaa gttggaacaa ttgaaaagcc 2520
ccaaataata aagcaaagca aaggggagaac tcaaaagcca ataaataatg gaggttacac 2580
ccagcaa 2587

```

<210> 105  
 <211> 1490  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3793128CB1

```

<400> 105
gtaagcccct tataaaacca catgtctcat ttgctgggtc caaatctctt ctttgtcctc 60
ttgaacctgg taacttcctt attgaggtta ataggggttc agcacaagag ctttaggagt 120
tatttggcta caccaggcc atttgctttt ctcaaggaag agataattgg cacactactc 180
ttaaatggca cctacactgc agtgggtgtg tatttttaca agggaagtca agccttcacc 240
tgtttccctc actttaactt acctgtgcc tgcagggtaa ttgtcagaga cttcagaaat 300
ccaagatcct ggggtccctt ttggacattg tgtcattgag gcatgtccct aggaacatgt 360
aagatgaagc catgataggc tattcaggac agttaccag aggaaatatt gattcacttt 420
ttatgcacaa catctgggta atggtttcta caatgagggt tatataattg gaatagggtga 480
ccagaaatth attgttaacc ttgtccgaag ttttaaagaa cactttcaca gctctcaaag 540
ctgaatgatg agtctcaact atttgactca cgtgaaatag aaaatctggt agcttttagac 600
tttctactag ccagtcatga tggatatctg gccattactg gcaatctgtg ttacacatga 660
ataaatacta ttagecaagt acaatgctct tctaaactga gggggtcagc cacttgactc 720
tccaaggaat atcctcctag attctagaca tattttcttg tcttgggttc agtaatttct 780
gtatatggct tggaggcatt ctgcaaagtt cagtcttcac tcttctactt tgtgtgcac 840
ttcaatcttc catgcttccg tgcagccact gtcacatcag atgatcgaaa atctcatccc 900
ttaacaaaat acacaaaata atagacactg atttaagttt agacactcta aaccttaaaa 960
aaaaaagatg ataatatcag ggctcatgac agtgatgtaa atctggaatg atactgcttt 1020
tgtggccaac cctttggcct ggcgaaaaga cggccaaaag ctggtgagac cagaactaga 1080
aagaccacct cctgttgtca cttgtgtgat tagaaaccaca aaatatctta tgaatttcat 1140
aacactttca tctttgtcga gaatctcacc tagttgcata agtctttaaa aaatcaacca 1200
gttcaaagat ttgctctcct tttatcccag tctattacct ttctgagttt aatccataag 1260
aaataaaaat ggtatatgca cttcctgtaa tgagatgcca atttagagtt gattccttat 1320
gttctctctt gccaaaagtaa gtgaatgaaa gccagttggc ttacatcata atagctttca 1380
tttaaggcac aggttttcaa ctttgtgtca ctgaaaatat cttaaagttaa tttacctgtg 1440
taattatcta aattatgctt ttcaagtttc tgtgcatctt cgggtgtcga 1490

```

<210> 106  
 <211> 1174  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4001243CB1

```

<400> 106
cgggacaaca ggaccctatg aagggtgggccc cacagcaaaa ggagagatga ttctagagca 60
tccagtcttc tagggcagca aaacaacctta aattttctaa gaggccaccc agctgagggt 120
gcccccgggg agggctgagg cgctcagggtg acggctccac tgcccactca cctgcgacct 180
caaagcccct ctctctcttg ggggtgctcct gacagccacc tccagggcag gcgagtggcg 240
ctgggacaaa ggctggcccc actgcgcccc acccaagcag acggctcctc ccccagacct 300
ggcgccaaac tggagtgaac gcccgaccac cgtgtctcac agggaaactg acaccagatg 360
cgaacttcca aatggatccc tccctgcaag tgtggagctg gcgctaccag gcaatgctct 420
ggccatgcgt ctaagacaca ggcagagggc gctgccacc acgctggcga cggcctcaaa 480
gcccctgttc atgcctggga cagcgcccaa ggaccttgct catgcctggg acaggcccca 540

```

```

gggccccac  tggctgcagt  cagcagcggg  cagggtggtg  gggaaggtg  tggacactcc  600
gtgggcccga  gctgggagaa  caaggcctat  tattggacac  ctggtggcca  tggcaaccac  660
acaaggatgc  ctgagactga  aaatctgttg  gcttcaagga  gctccagctc  ttgcaactggc  720
tgagtcacag  tgactatata  actcttactc  ccacttttgg  gacacttttt  gagagggaca  780
gggatcctat  ctaactacac  gggacagaca  tcgccccaga  ccgtcctgag  caagcctgga  840
cctgtgtgacc  ctaacgatga  aggtgtcccc  cagacaatgt  ccggggcagg  caccatgtctc  900
tcccaaccta  ccacagccag  atgtttttgt  aaagaacaat  aaaaatgaat  tactagaaaa  960
gcaaagactt  aaaatacaca  aaaaaaaaaa  caagggggag  gccgccgaat  atagaggacc  1020
cgggaagaccg  gggaattaat  cccgaaccgg  taccgtgggg  gcgctccaag  gattcccata  1080
tataggagc  cacattaaga  acttgggaaa  tcgagoccaa  tgcgtgacct  cgtgttgcca  1140
atgtaaccgg  acaattccca  caaccaaaac  aagg

```

&lt;210&gt; 107

&lt;211&gt; 818

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 6986717CB1

&lt;400&gt; 107

```

ctgggctcac  agacagggtg  tgagcaggca  gattcggggg  cagagagtgt  ggcaggacgc  60
tcagctctct  aatgacagcc  ctttcctggg  aacctcccc  tgtagcact  gccttactct  120
gtgggtctgt  ttggcctggg  agaggacagg  ccggaatgg  gtgggtgcgt  gcttctctat  180
aaaatgggaa  taaagacaat  atccatttca  catggcttct  gcgaggatga  aatggcacga  240
tatacgtaac  tcaactgtga  cttgggtcaa  tcaatgctgt  tttccctcct  ccgcttcctc  300
ttctatagat  ggtgattcca  ggattgacta  cattgctgat  aaaaactacc  ttctggggct  360
tccgttttgg  ggagctgggg  atggggagag  ggagtacaag  ttctagatgc  ctggtcagcc  420
cctctttttc  tcttctgcat  gtagggggg  gcttggacca  gcttgccctg  accctgcccc  480
aggagctgag  ggggaaggac  atgcggatgg  tccccatgga  gatgttcaac  tactgctccc  540
agctggagga  cgagaatagc  tcagctgggc  tggatattct  gggccaccct  gcaccaaggc  600
cagtcagag  cctgctaagc  ccaagcccgg  ggctgagccg  gagccggagc  ccagcacagc  660
ctgcccacag  aagcagaggg  accggccggc  gagcgtgagg  cgagccatgg  gcacggtgat  720
cattgcaggg  gtcgtgtgcg  gcgtcgtctg  catcatgatg  gtggtggccg  ctgcctatgg  780
ctgcatctac  gcctccctca  tggccaagta  ccaccgag

```

&lt;210&gt; 108

&lt;211&gt; 4717

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 7503512CB1

&lt;400&gt; 108

```

atggcgcggc  cgggtccggg  agggctcggg  gccccgcgcc  gctcgccttg  ccttctcctt  60
ctctggctgc  ttttgettcg  gctggagccg  gtgaccgcgc  cggccggccc  gcggggcgcc  120
tgcgcggccg  cctgcacttg  cgtcggggac  tcgctggact  gcggtgggcg  cgggctggct  180
gcgttgcccg  gggacctgcc  ctctcggagc  cggagcctaa  acctgagtta  caacaaactc  240
tctgagattg  accctgctgg  ttttgaggac  ttgccgaacc  tacaggaagt  gtacctcaat  300
aataatgagt  tgacagcggt  accatccctg  ggcgctgctt  catcacatgt  cgtctctctc  360
tttctgcagc  acaacaagat  tcgcagcgtg  gaggggagcc  agctgaaggc  ctacctttcc  420
ttagaagtgt  tagatctgag  tttgaacaac  atcacggaag  tgcggaacac  ctgctttcca  480
cacggaccgc  ctataaagga  gctcaacctg  gcaggcaatc  ggattggcac  cctggagttg  540
ggagcatttg  atggtctgtc  acggtcgttg  ctaactcttc  gcctgagcaa  aaacaggatt  600
cggctgatag  agggcctcac  cttccagggg  ctcaacagct  tggagggtgt  gaagcttcag  660
cgaaacaaca  tcagcaaact  gacagatggg  gccttctggg  gactgtccaa  gatgcatgtg  720
ctgcacctgg  agtacaacag  cctggtagaa  gtgaacagcg  gctcgtctta  cggcctcacg  780
gccctgcatc  agctccacct  cagcaacaat  tccatcgctc  gcattcaccg  caagggtggt  840
agcttctgcc  agaagctgca  tgagttggtc  ctgtccttca  acaacctgac  acggctggac  900
gaggagagcc  tggccgagct  gagcagcctg  agtgtcctgc  gtctcagcca  caattccatc  960

```

agccacattg	eggaggggtgc	cttcaagggga	ctcaggagacc	tgcgagtctt	ggatctggac	1020
cataacgaga	tttcgggcac	aatagaggac	acgagcggcg	ccttctcagg	gctcgacagc	1080
ctcagcaagc	tgactctgtt	tggaaacaag	atcaagtctg	tggctaagag	agcattctcg	1140
gggctggaag	gcctggagca	cctgaacctt	ggaggggaatg	cgatcagatc	tgtccagttt	1200
gatgcctttg	tgaagatgaa	gaatctttaa	gagctccata	tcagcagcga	cagcttcctg	1260
tgtgactgcc	agctgaagtg	gctgcccccg	tggctaattg	gcaggatgct	gcaggccttt	1320
gtgacagcca	cctgtgccca	cccagaatca	ctgaaggggtc	agagcatttt	ctctgtgccca	1380
ccagagagtt	tcgtgtgcga	tgacttctctg	aagccacaga	tcatacccca	gccagaaacc	1440
accatggcta	tgggtgggcaa	ggacatccgg	tttacetgct	cagcagccag	cagcagcagc	1500
tcccccatga	cctttgcctg	gaagaaagac	aatgaagtcc	tgaccaatgc	agacatggag	1560
aactttgtcc	acgtccacgc	gcaggacggg	gaagtgatgg	agtaaccac	catcctgcac	1620
ctccgtcagg	tcactttcgg	gcacgagggc	cgctaccaat	gtgtcatcac	caaccacttt	1680
ggctccacct	attcacataa	ggccaggctc	accgtgaatg	tgttgccatc	attcaccaaa	1740
acgccccacg	acataacccat	ccggaccacc	accatggccc	gcctcgaatg	tgctgccaca	1800
ggtaacccaa	accctcagat	tgcttgccag	aaggatggag	gcacggattt	ccccgctgcc	1860
cgtgagcgac	gcatacgtgt	catgcccggat	gacgacgtgt	ttttcatcac	tgatgtgaaa	1920
atagatgacg	caggggttta	cagctgtact	gctcagaact	cagccgggttc	tattttcagct	1980
aatgccaccc	tgactgtcct	agagacccca	tccttgggtg	tccccttgga	agaccgtgtg	2040
gtatctgtgg	gagaaacagt	ggccctccaa	tgcaaagcca	cggggaaccc	tcgcgcccg	2100
atcacctggg	tcaaggggga	ccgcccgtcg	agcctcactg	agcggcacca	cttgaccctt	2160
gacaaccagc	tcctggtggg	tcagaacgtg	gtggcagagg	atgcgggccg	atatacctgt	2220
gagatgtcca	acaccctggg	cacggagcga	gctcacagcc	agctgagcgt	cctgcccgcga	2280
gcaggctgca	ggaaggatgg	gaccacggta	ggcatcttca	ccattgctgt	cgtgagcagc	2340
atcgtcctga	cgtcactggg	ctgggtgtgc	atcatctacc	agaccaggaa	gaagagtcaa	2400
gagtacagtg	tcaccaacac	agatgaaacg	ctcgtgtccac	cagatgttcc	aagctacctc	2460
tcttctcagg	ggaccctttc	tgaccgacaa	gaaaccgtgg	tcaggaccga	gggtggccct	2520
caggccaatg	ggcacattga	gagcaatggg	gtgtgtccaa	gagatgcaag	ccactttcca	2580
gagcccgaca	ctcacagcgt	tgcttgccag	cagccaaagc	tctgtgctgg	gtctgcgtat	2640
cacaaagagc	cgtggaaaagc	gatggagaaa	gctgaaggga	cacctgggccc	acataagatg	2700
gaacacgggtg	gccgggtcgt	atgcagtgc	tgcaacaccg	aagtggactg	ttactccagg	2760
ggacaagcct	tccaccccca	gcctgtgtcc	agagacagcg	cacagccaag	tgccgcaaat	2820
ggcccggagc	cgggtgggag	tgaccaagag	cattctccac	atcaccagtg	cagcaggact	2880
gccgctgggt	cctgccccga	gtgccaaagg	tcgtcttacc	ccagtaacca	cgatagaatg	2940
ctgacggctg	tgaagaaaaa	gccaatggca	tctctagatg	ggaaagggga	ttcttccctg	3000
acttttagcaa	ggttgtatca	cccggactcc	acagagctac	agcctgcac	ttcattaact	3060
tcaggcagtc	cagagcgcgc	ggaagcccag	tacttgcttg	tttccaatgg	ccacctcccc	3120
aaagcatgtg	acgccagtc	cgagtccacg	ccactgacag	gacagctccc	cgggaaacag	3180
agggtgccac	tgctgttggc	accaaanaagc	taggttttgt	ctacctcagt	tcttgtcata	3240
ccaatctcta	cgggaaagag	aggtaggaga	ggctgcgagg	aagcttgggt	tcaagcgtca	3300
ctcatctgta	catagttgta	actcccatgt	ggagtatcag	tcgctcacag	gacttggatc	3360
tgaagcacag	taaacgcaag	aggggatttg	tgtacaaaag	gcaaaaaaag	tatttgatat	3420
cattgtacat	aagagttttc	agagatttca	tatatatctt	ttacagaggc	tatttttaac	3480
tttagtgcat	ggttaacaga	aaaaaattat	acaattttga	caatattatt	tttcgtatca	3540
gggtgctgtt	taattttgga	gggggtgggg	aaatagttct	ggtgccttaa	cgcatggctg	3600
gaatttatag	aggctacaac	cacattttgt	cacaggagtt	tttgggtcgg	gggtgggaagg	3660
atggaaggcc	ttggatttat	attgcacttc	atagaccctt	aggctgctgt	gcggtgggac	3720
tccacatgcg	ccggaaggag	cttcagggtga	gcactgctca	tgtgtggatg	ccctgcaac	3780
aggcttccct	gtctgtagag	ccaggggtgc	aagtgccatc	cacacttgca	gtgaatggct	3840
tttccctttta	ggtttaagtc	ctgtctgtct	gtaaggcgta	gaatctgtcc	gtctgtgaag	3900
cgtagaatga	gggttggtta	tccatcacaa	gcaaaaggct	agaacagtta	aacactgcct	3960
ttcctcctcc	tcttatttta	tgataaaagc	aaatgtggcc	ttctcagtat	cattcgattg	4020
ctatttgaga	cttttaaaatt	aaggtaaagg	ctgctggtgt	tggtacctgt	ggatttttct	4080
atactgatgt	tttcgttttg	ccaatataat	gagtattaca	ttggccttgg	gggacagaaa	4140
ggaggaagtt	ctgacttttc	agggctacct	tatttctact	aaggacccag	agcaggcctg	4200
tccatgccat	tccttcgcac	agatgaaact	gagctgggac	tggaaggagc	agcccttgac	4260
ctgggttctg	ggtataattt	gcacttttga	gactggtagc	taaccatctt	atgagtgccca	4320
atgtgtcatt	tagtaaaact	taaatagaaa	caaggctcct	caaagtgtcc	tttggccaaa	4380
agctgaaggg	agttactgag	aaaatagtta	acaattactg	tcagggtgtca	tcactgttca	4440
aaaggtaagc	acatttagaa	ttttgttctt	gacagttaac	tgactaatct	tacttccaca	4500
aaatatgtga	atttgctgct	tctgagaggc	aatgtgaaa	agggagtatt	acttttatgt	4560
acaaagttaa	ttatttatag	aaatttttgt	acagtgtaca	ttgaaaacca	tgtaaaatat	4620
tgaagtgtct	aacaaatggc	attgaagtgt	ctttaataaa	ggttcattta	taaatgtcaa	4680
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	atcggctc			4717